Decision support model for the patient admission scheduling problem with random arrivals and departures

Aregawi K. Abera ∗
University of Tasmania
TAS 7001, Australia
aregawi.abera@utas.edu.au

Malgorzata M. O’Reilly ∗†
University of Tasmania
TAS 7001, Australia
malgorzata.oreilly@utas.edu.au

Barbara R. Holland
University of Tasmania
TAS 7001, Australia
barbara.holland@utas.edu.au

Mark Fackrell ∗†
The University of Melbourne
VIC 3010, Australia
fackrell@unimelb.edu.au

Mojtaba Heydar ∗
University of Tasmania
TAS 7001, Australia
mojtaba.heydar@utas.edu.au

ABSTRACT

Introduction. The patient admission scheduling (PAS) problem is a class of scheduling problems that must be handled by the managers of the hospital admission systems. The problem arises when patients arriving at the hospital need to be allocated to beds in an optimal manner, subject to the availability of beds and the needs of patients.

The PAS problem in a dynamic context, as analysed in Ceschia and Schaerf [2] and Lusby et al. [6], considers a scenario in which random arrivals and unknown departures of patients are gradually revealed over the planning horizon. The problem was formulated as an integer programming model, and various procedures for computing the optimal solution were proposed. Ceschia and Schaerf [2] developed a metaheuristic algorithm based on simulated annealing and neighborhood search. Lusby et al. [6] developed an adaptive large neighbourhood search procedure to solve the problem.

Although the arrivals and departures of patients are in general random, the models in [2, 6] assumed deterministic inputs such as a fixed length of stay for each patient, and a fixed number of arrivals at the start of each day. Here, we build on the analysis in Lusby et al. [6], and develop a model for the PAS problem in a dynamic context, which captures the random dynamics of the flow of the patients.

Our aim here is to develop an improved mathematical model to solve the PAS problem in a dynamic environment with random arrivals and departures. At the start of each day we record new information about the registered patients, newly arrived patients and future arrivals (including emergency patients and scheduled arrivals), and then determine an optimal assignment of patients to beds. Our goal is to provide a decision support tool for the patient scheduling process to be used by hospital administrators and planners.

Notation. We use similar notation to Demeester et al. [3] and Turhan and Bilgen [8] for the parameters and variables of our model, with some minor changes.

- Patients are classified into three groups, admitted patients, planned patients, and emergency patients. Admitted patients are patients that are successfully admitted to the hospital, and allocated to a bed. Planned patients have not been admitted to the hospital as yet, but have pre-determined admission dates, denoted by $d_{plan}^p$. Emergency patients are patients whose admission date is equal to their registration date, that is, $d_{plan}^p = d_{reg}^p$, since their admission cannot be postponed and is unplanned.

- Patients are denoted by $p$, with $p \in P$, where $P$ is the set of all patients. Also let $M \subset P$ be the set of all male patients, and $F \subset P$ be the set of all female patients. Patients have the following properties: admission date and a discharge date, age and gender, required treatment, and room preference.

- Days are denoted by $d$, with $d \in D$, where $D = \{0, 1, \ldots, D\}$ is the set of all days in the planning period of the time horizon. Further, let $d_p \in D$ be the admission day of patient $p$, where $D_p = \{d_{plan}^p, \ldots, d_{max}^p\} \subseteq D$ is the set of acceptable days for patient $p$ to be admitted to the hospital.

- The length of stay (LoS) of patient $p$ is denoted by $L_p$. This is a random variable recording the number of days patient $p$ will stay in the hospital till he/she gets discharged. We assume $L_p$ takes values $l_p = 0, 1, \ldots, l_{max}^p$, for some positive integer $l_{max}^p$.

- A hospital consists of different wards. Typically, each ward is specialized in treating one kind of pathology...
such as cardiovascular diseases, oncology, or dermatology, which is considered as the major specialism of the ward as in Demeester et al. [3]. Wards can also perform other treatments as minor specialisms. Wards are denoted by \( W_i, i = 1, \ldots, W \), where \( W \) is the total number of wards in the hospital. Wards can support one or more specialisms \( S_u, u = 1, \ldots, S \), where \( S \) is the total number of specialisms. We write
\[
S_u \sim W_i \tag{1}
\]
when specialism \( S_u \) is available in ward \( W_i \), and
\[
S(p) = S_u \tag{2}
\]
when patient \( p \) requires specialism \( S_u \). Patients admitted to ward \( W_i \) may have to be in a particular age range, between some minimum \( a(W_i) \) and maximum \( A(W_i) \).

- Rooms are denoted by \( r \), with \( r \in R = \{1, \ldots, R\} \), where \( R \) is the total number of rooms in the hospital. We write
\[
r \in W_i \tag{3}
\]
when room \( r \) is in ward \( W_i \subset R \). A room can be described by its age policy, gender policy and by its special features, such as the presence of oxygen, nitrogen, telemetry or television. A room may support one or more different specialisms \( S_u \), depending on the room features. Rooms have a specified gender policy, which is one of the following: male only \( M \), female only \( F \), depends on the gender of the first patient \( SG \) (same-gender policy), or all genders are allowed \( N \). It is preferable not to assign male and female patients to the same room at the same time. We denote by \( R^M, R^F, R^{SG}, R^N \subset R \) the sets of all rooms with policies \( M, F, SG, N \), respectively.

- The capacity of room \( r \) is denoted by \( \kappa_r \). This is the total number of beds in room \( r \).

- Assignment \( \sigma \) is the collection of decisions \( x_{p,r,d}(\sigma) \) and \( y_{p,d}(\sigma) \) defined as,
\[
x_{p,r,d}(\sigma) = \begin{cases} 1 & \text{if patient } p \text{ is assigned to room } r \text{ on day } d \\ 0 & \text{otherwise}, \end{cases} \tag{4}
\]
\[
y_{p,d}(\sigma) = \begin{cases} 1 & \text{if patient } p \text{ is admitted on day } d \\ 0 & \text{otherwise}, \end{cases} \tag{5}
\]
and note that \( y_{p,d}(\sigma) = 1\{d = d_p(\sigma)\} \), where \( 1\{\cdot\} \) is an indicator function.

- In order to calculate the violation of gender policy, we define the presence of male, female and both patients in room \( r \) on day \( d \) as follows,
\[
m_{r,d}(\sigma) = \begin{cases} 1 & \text{if there is at least one male patient in room } r \text{ on day } d \\ 0 & \text{otherwise}, \end{cases} \tag{6}
\]
\[
f_{r,d}(\sigma) = \begin{cases} 1 & \text{if there is at least one female patient in room } r \text{ on day } d \\ 0 & \text{otherwise}, \end{cases} \tag{7}
\]
\[
b_{r,d}(\sigma) = \begin{cases} 1 & \text{if both genders are present in room } r \text{ on day } d \\ 0 & \text{otherwise}. \end{cases} \tag{8}
\]

- Given patient \( p \), the required features of a room for allocation are grouped into two categories, needed room feature (NRF), and preferred room features (PRF).

Given feature \( j \) of some room \( r \), we write
\[
\text{NRF}_j(p,r)(\sigma) = \begin{cases} 1 & \text{if the needed room feature is provided} \\ 0 & \text{otherwise}. \end{cases} \tag{9}
\]
Similarly, we write
\[
\text{PRF}_j(p,r)(\sigma) = \begin{cases} 1 & \text{if the preferred room feature is provided} \\ 0 & \text{otherwise}. \end{cases} \tag{10}
\]

- Transfer means relocating a patient from one room to another during their stay. As it is described by Demeester et al. [3], transfers can be planned or unplanned, the latter should be avoided if possible. As an example of a planned transfer, a patient might be transferred from surgery to an intensive care unit, and after recovery they might be transferred to another ward. An unplanned transfer could be due to a shortage of resources such as beds or rooms. The transfer of patient \( p \) from room \( r \) to another room \( r^* \) on day \( d \) is recorded using variable
\[
t_{p,r,r^*,d}(\sigma) = 1\{x_{p,r^*,d-1}(\sigma) = 1, x_{p,r^*,d}(\sigma) = 1, r^* \neq r\}. \tag{11}
\]
That is, \( t_{p,r,r^*,d}(\sigma) = 1 \) when patient \( p \) was transferred from room \( r \) to room \( r^* \neq r \) on day \( d \), and \( t_{p,r,r^*,d}(\sigma) = 0 \) otherwise.

- Let \( Q_{r,d}(\sigma) \) be the event that a gender conflict is observed in room \( r \) on day \( d \), given assignment \( \sigma \). Also, define the random variable \( b_{r,d}(\sigma) \) such that \( b_{r,d}(\sigma) = 1 \) if the event \( Q_{r,d}(\sigma) \) occurs, and \( b_{r,d}(\sigma) = 0 \) otherwise. That is,
\[
b_{r,d}(\sigma) = 1\{Q_{r,d}(\sigma)\}. \tag{12}
\]
Then the mean value of \( b_{r,d}(\sigma) \) is equal to the probability of the event \( Q_{r,d}(\sigma) \) occurring, with
\[
E(b_{r,d}(\sigma)) = P_r(Q_{r,d}(\sigma)). \tag{13}
\]
Denote by \( A_{m,d}(\sigma) \) and \( A_{f,d}(\sigma) \) the events that all males have left the room before day \( d \), and that all females have left the room before day \( d \), respectively. \( F_{r,d} \) is the set of all female patients assigned to room \( r \) on day \( d \), that is \( F_{r,d} = \{p \in F : x_{p,r,d}(\sigma) = 1\} \), and \( M_{r,d} \) is the set of all male patients assigned to room \( r \) on day \( d \), which is \( M_{r,d} = \{p \in M : x_{p,r,d}(\sigma) = 1\} \). Then,
\[
1 - P_r(Q_{r,d}(\sigma)) = P_r(A_{m,d}(\sigma)) + P_r(A_{f,d}(\sigma)) - P_r(A_{m,d}(\sigma) \cap A_{f,d}(\sigma)) = \prod_{M_{r,d}} x_{p,r,d}(\sigma)P_r(L_p < d - d_p(\sigma)) + \prod_{F_{r,d}} x_{p,r,d}(\sigma)P_r(L_p < d - d_p(\sigma)) - \prod_{M_{r,d} \cap F_{r,d}} x_{p,r,d}(\sigma)P_r(L_p < d - d_p(\sigma)). \tag{14}
\]
• Let $Z_{p,r,d}(\sigma)$ be a random variable such that, given assignment $\sigma$, $Z_{p,r,d}(\sigma) = 1$ if patient $p$ is in room $r$ on day $d$, and $Z_{p,r,d}(\sigma) = 0$ otherwise.

• Let $Y_{r,d}(\sigma) = \sum_{p\in P} Z_{p,r,d}(\sigma)$ be a random variable recording the number of patients in room $r$ on day $d$, given assignment $\sigma$, and $(E(Y_{r,d}(\sigma)) - \kappa_r)$ be the expected excess in room $r$ on day $d$. We then have

$$E(Y_{r,d}(\sigma)) = E\left(\sum_{p\in P} Z_{p,r,d}(\sigma)\right) = \sum_{p\in P} E(Z_{p,r,d}(\sigma)) = \sum_{p\in P} P_r(Z_{p,r,d}(\sigma) = 1) = \sum_{p\in P} x_{p,r,d}(\sigma) P_r(L_p \geq d - d_p(\sigma)).$$

(15)

• We define the following cost functions, which we later use as coefficients in the objective function. Let $c_{p,r,d}$ be the cost of assigning patient $p$ to a room $r$ on day $d$. Let $c_{p,r,r',d}^{(T)}$ be the cost of transferring patient $p$ from room $r$ to room $r'$ on day $d$, with $c_{p,r,r',d}^{(T)} = 0$. Let $c_{r,d}^{(G)}$ be the penalty incurred for the violation of gender policy in room $r$ on day $d$. Let $c_{r,d}^{(O)}$ be the penalty incurred when the capacity $\kappa_r$ of room $r$ is exceeded on day $d$. Let $c_{r,d}^{(D)}$ be the penalty incurred for the admission delay of patient $p$ on day $d$.

Using the parameters and variables mentioned above, we now construct a stochastic integer programming model with suitable constraints due to patients medical needs and age, room capacity, and gender policy, similar to Lushby et al. [6], with suitable modifications. These include hard constraints that must be met and soft constraints that can be violated when necessary, but which are subject to cost penalties.

**Hard constraints.** For a solution to be feasible, it has to satisfy the following set of hard constraints (16)-(21). First, we set the room capacity constraints,

$$\sum_{p\in P} x_{p,r,d}(\sigma) \leq \bar{\kappa}_r, \quad \forall r \in R, \forall d \in D,$$

(16)

where $\bar{\kappa}_r \geq \kappa_r$ is some maximum allowed threshold for the total number of patients in room $r$, after taking into account an overstay risk.

Next, patient $p$ should be assigned to ward $W_i$ that is suited for the patient’s age, denoted $A_p$. The minimum age limit $a(W_i)$ and maximum age limit $A(W_i)$ allowed in ward $W_i$ should be respected. Therefore,

$$x_{p,r,d}(\sigma) 1\{r \in W_i\} \leq 1\{a(W_i) \leq A_p \leq A(W_i)\}, \quad \forall p \in P, \forall r \in R, \forall d \in D.$$  

(17)

Furthermore, a patient $p$ should be assigned to a ward $W_i$ with a suitable specialization $S_u$, for some $u$. Therefore,

$$x_{p,r,d}(\sigma) 1\{r \in W_i, S(p) = S_u\} \leq 1\{S_u \sim W_i\}, \quad \forall p \in P, \forall r \in R, \forall d \in D.$$  

(18)

Additionally, the medical treatment of a patient $p$ may require that he/she is assigned to a room $r$ with special equipment or other features required for the treatment. That is, when making decision $x_{p,r,d}(\sigma) = 1$ we must have $r$ such that $NRF_j(p, r) = 1$, when patient $p$ requires room feature $j$. Therefore,

$$x_{p,r,d}(\sigma) \leq 1\{NRF_j(p, r) = 1\}, \quad \forall p \in P, \forall r \in R, \forall d \in D.$$  

(19)

Also, patients have to be admitted within the planning horizon, and so

$$\sum_{d \in D} y_{p,d}(\sigma) = 1, \quad \forall p \in P.$$  

(20)

Moreover, if patient $p$ is admitted on day $\bar{d}$, the patient must appear in some room $r$ the following $t_p^{max} - 1$ nights, which gives,

$$\sum_{r \in R} x_{p,r,d}(\sigma) \geq y_{p,d}(\sigma), \quad \forall p \in P, \forall d = \bar{d}, \ldots, \bar{d} + t_p^{max} - 1, \forall \bar{d} \in D_p.$$  

(21)

**Soft constraints.** The set of soft constraints (22)-(25) corresponds to desirable conditions that do not have to be met, but are subject to penalties.

Ideally, patients should be allocated as per their gender to an appropriate room $r$ with its specified gender policy. We evaluate the presence of a female patient $f_{r,d}(\sigma)$ in room $r$ on day $d$ is using

$$f_{r,d}(\sigma) \geq x_{p,r,d}(\sigma), \quad \forall p \in F, \forall r \in R^{SG}, \forall d \in D.$$  

(22)

and the presence of a male patient $m_{r,d}(\sigma)$ using

$$m_{r,d}(\sigma) \geq x_{p,r,d}(\sigma), \quad \forall p \in M, \forall r \in R^{SG}, \forall d \in D.$$  

(23)

To determine when both genders are present $b_{r,d}(\sigma)$ we use the following constraint,

$$b_{r,d}(\sigma) \geq m_{r,d}(\sigma) + f_{r,d}(\sigma) - 1, \quad \forall r \in R^{SG}, \forall d \in D.$$  

(24)

The transfer of patients is handled using the following constraint,

$$t_{p,r,r^{*},d}(\sigma) \geq x_{p,r,d}(\sigma) - x_{p,r,d-1}(\sigma), \quad \forall p \in P, \forall r \in R, \forall d = 2, \ldots, D.$$  

(25)

Some other desirable conditions could also be considered. A patient who asked for a single room, in case of lack of single rooms should preferably be assigned to a twin room.

In addition to major medical treatment, a patient $p$ may need to undergo other minor medical treatments within department $W_i$ in a room $r$ with special equipment to treat the patient assigned, which requires some minor specialization $S_r$ for some suitable $\ell$.  

12
A patient $p$ may prefer a room $r$ with features that in some degree correspond to the specialism that is required to treat the patient’s clinical condition. That is, it is preferable to have $\text{PREF}_j(p, r) = 1$, when patient $p$ prefers room feature $j$.

**Objective function.** We define the stochastic objective function as the total expected cost incurred over the planning horizon $\mathcal{D} = \{0, 1, \ldots, D\}$, and write it as a sum of the following cost components. The first component captures the cost of assigning patient $p \in \mathcal{P}$ to room $r \in \mathcal{R}$ on day $d \in \mathcal{D}$. The second component calculates the cost of transferring patient $p \in \mathcal{P}$ from room $r \in \mathcal{R}$ to another room $r' \in \mathcal{R}$ on day $d \in \mathcal{D}$. The third component determines the penalty incurred for the violation of gender policy in room $r \in \mathcal{R}$ on day $d \in \mathcal{D}$. The fourth component determines the penalty incurred when the capacity $\kappa_r$ of room $r \in \mathcal{R}$ is exceeded on day $d \in \mathcal{D}$. The fifth component computes the penalty incurred when the admission of patient $p \in \mathcal{P}$ is delayed beyond the maximum acceptable admission day $d_p^{\text{max}}$ on day $d \in \mathcal{D}$. The resulting expression is stated as,

$$
\begin{align*}
\min_{\sigma} \left\{ & \sum_{p \in \mathcal{P}} \sum_{d \in \mathcal{D}} \sum_{r \in \mathcal{R}} c_{p, r, d} \times x_{p, r, d}(\sigma) \times \Pr(L_p \geq d - d_p(\sigma)) \\
& + \sum_{p \in \mathcal{P}} \sum_{d \in \mathcal{D}} \sum_{r \in \mathcal{R}} c_{p, r, r', d}^{(T)} \times y_{p, r, r', d}(\sigma) \times \Pr(L_{p} \geq d - d_p(\sigma)) \\
& + \sum_{d \in \mathcal{D}} \sum_{r \in \mathcal{R}} c_{r, d}^{(G)} \times \Pr(Q_{r, d}(\sigma)) \\
& + \sum_{d \in \mathcal{D}} \sum_{r \in \mathcal{R}} c_{r, d}^{(D)} \times \sum_{d \in \mathcal{D}} \left( \frac{\max\{0, E(Y_{r, d}(\sigma)) - \kappa_r\}}{\kappa_r} \right) \\
& + \sum_{p \in \mathcal{P}} c_{p, d}^{(P)} \times \sum_{d \in \mathcal{D}} \left( \frac{d - d_{p}^{\text{plan}}}{d_{p}^{\text{max}} - d_{p}^{\text{plan}}} \right) \times y_{p, d}(\sigma) \right\}, \quad (26)
\end{align*}
$$

where $\max\{0, E(Y_{r, d}(\sigma)) - \kappa_r\}$ is the expected number of patients in room $r$ on day $d$ above the capacity of room $r$, given assignment $\sigma$.

**Random arrivals and departures.** In order to model the random departures, we assume that the random variable $L_{p}$ that records the LoS of the type-$p$ patient, and takes values $\ell_{p} = 0, 1, \ldots, \ell_{p}^{\text{max}}$, for some positive integer $\ell_{p}^{\text{max}}$, follows a discrete phase-type distribution in Latouche and Ramaswami [5, Chapter 2] and Neuts [7] with parameters that depend on $p$.

That is, we consider a discrete-time Markov chain with state space $\mathcal{V} = \{0, 1, \ldots, \ell_{p}^{\text{max}}\}$, where $\ell_{p}^{\text{max}}$ is an absorbing state, and one-step transition probability matrix $\mathbf{P}$ given by

$$
\mathbf{P}^* = \begin{bmatrix}
\mathbf{P} & \mathbf{p} \\
0 & 1
\end{bmatrix}, \quad (27)
$$

for some matrix $\mathbf{P} = [P_{i,j}], i,j=0,1,\ldots,\ell_{p}^{\text{max}}-1$ and (column) vector $\mathbf{p} = [p_{\ell_{p}^{\text{max}}}, \ldots, p_{0}]$, and the initial distribution (row) vector $\tau = [\tau_{i}]_{i=0,1,\ldots,\ell_{p}^{\text{max}}-1}$.

We then assume that the random variable $L_{p}$ follows discrete phase-type distribution with parameters $\tau$ and $\mathbf{P}$, which models time till absorption in the above chain,

$$
L_{p} \sim \text{PH}(\tau, \mathbf{P}), \quad (28)
$$

which gives, for $\ell_{p} = 0, 1, \ldots, \ell_{p}^{\text{max}}$,

$$
\begin{align*}
\Pr(L_p = \ell_p) &= \tau \mathbf{P}^\ell \mathbf{p}, \quad (29) \\
\Pr(L_p \leq \ell_p) &= 1 - \tau \mathbf{P}^\ell \mathbf{1}, \quad (30)
\end{align*}
$$

where $\mathbf{1}$ is a (column) vector of ones of appropriate size.

We use these expressions in order to evaluate the first two components of the objective function in (26).

In order to include the random arrivals that may occur during the planning horizon, we apply an approach similar to Kumar et al. [4]. We simulate random arrivals (from a suitable distribution) multiple times, resulting in a number of possible solutions. We then compare the different solutions by running simulations over some long time period, and then choose the preferred solution.

For example, suppose that the arrivals of patients (emergency or scheduled) occur according to a Poisson process with rate $\lambda_p$ per day, for type-$p$ patient, for all $p \in \mathcal{P}$, where patient type is determined by their medical needs, age and gender. We generate the random arrivals of emergency patients in the time horizon $[0, D]$, using standard simulation methods. As one possibility, for each patient type $p$, assume that $[D \lambda_p]$ arrivals have occurred during the time interval $[0, D]$, and then draw the random arrival times from a discrete uniform distribution on $[0, 1, \ldots, D]$. We then add the set of such generated patients to the problem, and solve it using the model in (26), treating these patients as registered patients, and so, patients that are known to the system.

**Solution approach.** We use simulation in order to generate random inputs for our model, and apply metaheuristic algorithms, similar to [6], including greedy search, adaptive neighbourhood search and simulated annealing, to solve the stochastic integer program. In our algorithm, we set the initial solution to be the optimal solution of the algorithm in [6], and compare our results with those of Lusby et al. [6]. The results of the application of our model will be reported in [1].

### 1. REFERENCES


