Report of Simulation Model for Hospital Outpatiant and Laboratory Stations

August 11, 2021

Contents

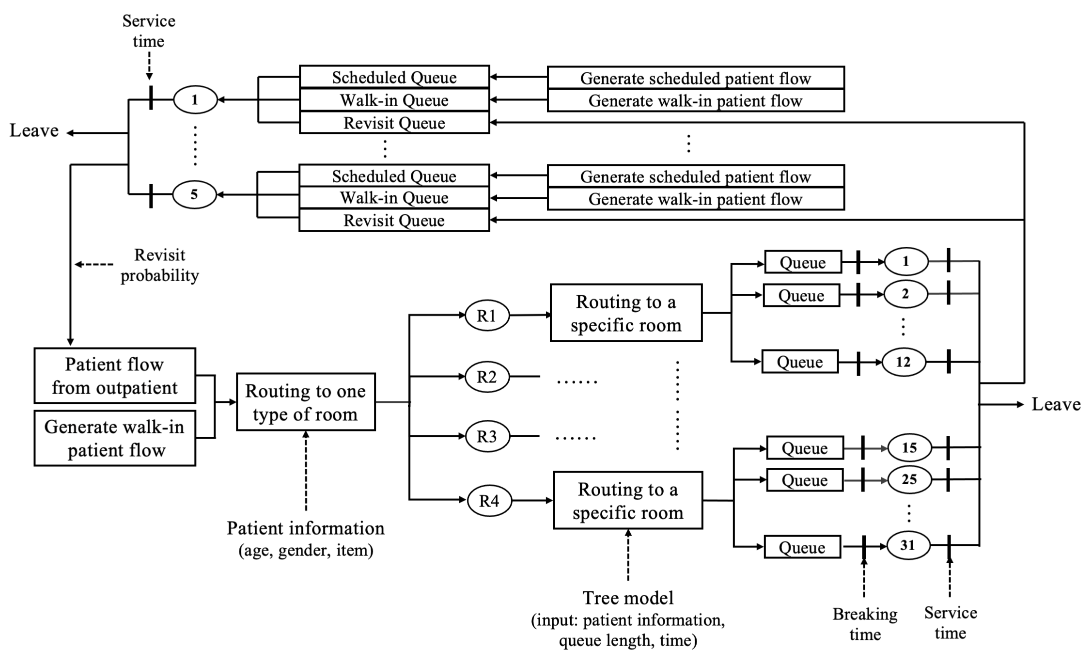
1 Introduction

1.1 Parameter Setting

Parameter	Meaning	Value
T	when the laboratory closes	50
$\overline{T}$	when the outpatient stops admitting new walk-ins	70
T_{max}	when there is no more service	130
Lunch_s	when lunch break starts	41
Lunch_e	when lunch break ends	60
c_w	unit cost of a walk-in's waiting per slot	1
c_{comp}	unit cost of a doctor's working per slot	20
$\overline{\gamma}$	schduled patients' show up probability	0.8
M	number of scheduled patients to schedule	21
K	number of scenarios	100

Table 1: Parameter Setting. Each unit time (slot) represents 6 minutes. T, $\overline{T}$, and T_{max} are counting the slots since 8am excluding the lunch break time.

1.2 Patient Flow



1.3 Outpatient Station

At the outpatient station, there are 5 identical doctors, sharing the same distribution of scheduled and walkin patients arrival, giving the same priority to serve different kinds of patients, and sharing the same distribution of service time.

1.3.1 Arrival

For the jth doctor, there are three kinds of patients – scheduled patients, walkin patients and revisit patients, and we denote the ith arrival of each of them as $p_i^{j,s}$, $p_i^{j,w}$, and $p_i^{j,r}$, respectively. Upon their arrival, they join the queue $q^{j,s}$, $q^{j,w}$, and $q^{j,r}$ correspondingly. Among all the patients, their **arrival** is as follows.

- Firstly, the arrvial of scheduled patients (ex) is determined by appointment schedule (AT) $\boldsymbol{z} = (z_1, \ldots, z_{\bar{T}})$, $z_t \in \{0, 1\}$ for $t = 1, \ldots, \bar{T}$ and show-up probability $\bar{\gamma}$. Whether a scheduled patient shows up or not follows a Bernoulli distributuon, and we use γ_t for $t = 1, \ldots, \bar{T}$ to indicate if a scheduled patient scheduled at slot t shows up ($\gamma_t = 1$) or not ($\gamma_t = 0$). Therefore, there is a **scheduled** patient arriving at the beginning of slot t if $z_t = 1$ and $\gamma_t = 1$.

- Secondly, the arrival of walkin patients (ex) follows time-nonhomogeneous Poisson. The number of **walkin** patients arriving at the beginning of each slot is denoted by $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_{\bar{T}})$, for $t = 1, \ldots, \bar{T}$.

- Thirdly, the arrival of revisit patients (en) is determined by reentry probability (p_1 for scheduled and walkin patients, p_2 for first-time revisit patients, and 0 for other patients), walking time, sojourn time at the lab test station, and reentry schedule (RT) $\boldsymbol{r} = (r_1, \ldots, r_T)$, $r_t \in [1, \bar{T}]$ for $t = 1, \ldots, T$. Whether a patient served at slot t needs a revisit or not follows a Bernoulli distribution, and we use η_t to indicate if he needs a revisit ($\eta_t = 1$) or not ($\eta_t = 0$). The walking time is sampled from a time-varying distribution of walking time between the outpatient station and the lab test station, and it will be sampled twice. The back time for a revisit patient generated at slot t is denoted by Δ_t. Therefore, the arrival time for a **revisit** patient generated at slot t is $max\{\Delta_t, r_t\}$ if $\eta_t = 1$.

1.3.2 Service

The **priority** of the three queues are $q^{j,s} > q^{j,r} > q^{j,w}$, meaning that the doctor will serving the queue with higher priority whenever he is available and the queue is nonempty.

The **service time** for each patient (ex) can either be determined as the length of a slot ($6mins$) or be random and follow a lognormal distribution.

1.4 Laboratory Station

1.4.1 Arrival

There are two kinds of patients at the laboratory station, walkin patients and patients from the outpatient station generated by the 5 doctors described above, and we denote the *ith* arrival of the former kind as p_i. We further denote the collection of all patients visiting the lab station as P^1 and P^2 for the above two kinds respectively. Among all these patients, they arrive with their own patient information (ex), specifically, their lab test item, age range, and gender. We denote the collection of patient information as $INFO = \{(item, age, gender) : item \in \{A, \dots, F\}, age \in \{3, \dots, 9\}, gender \in \{F, M\}\}$.

- The arrival of **walkin** patients at the laboratory station (ex) follows time-nonhomogeneous Poisson distribution.

  For each patient p_i, we will assign him with some patient information $\boldsymbol{f^{p_i}} \in INFO$ (ex) according to the distribution F^w of walkin patients' information.

- The arrival of **patients from the outpatient station** (en) is determined by $p_i^{j,s}$, $p_i^{j,w}$, and $p_i^{j,r}$, $j \in \{1, 2, 3, 4, 5\}, i \in \{0, 1, 2, \dots\}$ and reentry probability as described above.

  For each patient among p them, we will first assign him with a patient type $e \in TYPE = \{1, 2, 3\}$ (ex) according to the time-varying distribution T of patient type. Then, we assign him with some patient information $\boldsymbol{f^p} \in INFO$ (ex) according to the distribution F^e of outpatients' information, and the distribution is determined by his patient type e.

1.4.2 Routing

Upon patients' arrival, they immediently join the same queue for routing. There are two level routing – the first one routing to a room type and the second one routing to a specific room. Note that there is one doctor in each room, so we may use room and doctor interchangeably.

- The first level of routing is to assign a patient to one of four room types, which we denote as $RMTP = \{1, 2, 3, 4\}$.

  Firstly, the probability of routing to each type $P = \{(P_1, P_2, P_3, P_4) : P_i \in [0, 1], i = 1, 2, 3, 4\}$ is decided by a tree model. There are six features for the tree model and we denote the collection of features for the tree model as $F = \{(a, it, h, w, \boldsymbol{ql}, num)\}$, whose meanings are summarized in Table 2. We use a function $M : F \to P$ to denote such tree model. Then for each patient p, we can generate a feature vector $f^p \in F$ for him upon his arrival and obtain $M(f^p)$, which is a distribution of room type.

  Secondly, we sample from $M(f^p)$ to obatin his room type rt^p (en).

- The second level of routing is to assign patient p to one of the rooms of his room type rt^p. Similar to the first level decision, we use tree model $M^{rt^p} : F \to P, rt^p \in \{1, 2, 3, 4\}$

Notation	Feature Name	Explanation
a	age range	age range of patient
it	item type	classified examination item types
h	arrival hour	hour of the day when patient arrives
w	weekday	day of the week when patient arrives
$\boldsymbol{ql}$	queue length	the number of patient in queues when patient arrives
num	number of Server	number of open test rooms when patient arrives

Table 2: Features for routing model. In the first level of routing, $\boldsymbol{ql}$ contains the total queue length of each room type, and in the second level of routing, $\boldsymbol{ql}$ contains the queue length of each room within some room type determined by the first level of routing.

to decide the probability of routing to each room, and sample from $M^{rt^p}(f^p)$ to obatin his room r^p (en).

Note that the service time for routing is 0, so patient p immediently join the queue of room r^p upon his arrival.

1.4.3 Service

For each tuple of (ultrasonic item, room type, time), we have a distribution of lab service time. For a patient p entering the lab service at time t^p, we can generate a tuple (f_1^p, rt^p, t^p) for him. Then the service time for him (en) is sampled from the corresponding distribution determined by (f_1^p, rt^p, t^p).

2 Implementation Details

2.1 Big Iteration Algorithm

The big iteration algorithm is as follows:

where the Algorithm 1 is:

and function p is given as:

$$p(\boldsymbol{l})_t = \begin{cases} p_1 & l_t \text{ serves a scheduled or walkin patient} \\ p_2 & l_t \text{ serves a first-time reentrant patient} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Note that in the big iteration algorithm, Algorithm 1 is applied on an approximation model of the outpatient station, with the following assumptions:

- The **service time** for each patient is fixed and equals the length of a slot (6 minutes). As a result, the patients arrival at the lab test station can only by a multiple of 6.

Data: $\boldsymbol{y}^0, \boldsymbol{\beta}, \boldsymbol{\Delta}^0$
Result: $\boldsymbol{z}^{GRD}$, $\boldsymbol{r}^{GRD}$,$\boldsymbol{y}^{GRD}$
Initialize: $\boldsymbol{z}^{GRD} = \emptyset$, $\boldsymbol{r}^{GRD} = \emptyset$;
for $k = 0, 1, 2, \ldots, \overline{T}$ **do**
  Apply Algorithm 1 using $\boldsymbol{y}^k$ and $\boldsymbol{\Delta}^k$, and get $\boldsymbol{z}^{k+1}$, $\boldsymbol{r}^{k+1}$;
  Compute y^{k+1} using Equation $\boldsymbol{y}^{k+1} = \mathbb{E}[p(\boldsymbol{\ell}(\boldsymbol{\eta}, \boldsymbol{\Delta}, \boldsymbol{\beta}|\boldsymbol{y}^k, \boldsymbol{z}^{k+1}, \boldsymbol{r}^{k+1}))] := g(\boldsymbol{y}^k)$;
  Generate arrival at the laboratory test station by $\boldsymbol{y}^{k+1}$, and use simulation of the laboratory test station to get the new empirical distribution $\boldsymbol{\Delta}^{k+1}$;
end
$\boldsymbol{r}^{GRD} \leftarrow \boldsymbol{r}^{\overline{T}+1}$;
$\boldsymbol{y}^{GRD} \leftarrow \boldsymbol{y}^{\overline{T}+1}$;

Algorithm 3: Iterative Greedy Algorithm to solve $(\boldsymbol{z}^{GRD}, \boldsymbol{r}^{GRD}, \boldsymbol{y}^{GRD})$

Data: $\boldsymbol{y}^k, \boldsymbol{\beta}, \boldsymbol{\Delta}$
Result: $\boldsymbol{z} := (z_1, z_2, \ldots, z_M)$, $\boldsymbol{r} := (r_1, r_2, \ldots, r_T)$
Initialize: $\boldsymbol{z} = \emptyset$, $\boldsymbol{r} = \emptyset$;
Generate K scenarios of $(\boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Delta})$;
for $m = 1, 2, \ldots, M$ **do**
  for $t = 1, \ldots, \overline{T}$ **do**
    $a_t \leftarrow V(\boldsymbol{z}\perp\{t\}, \boldsymbol{0}) - V(\boldsymbol{z}, \boldsymbol{0})$;
  end
  $z_m \leftarrow \text{argmax}\{a_t | t = 1, 2, \ldots, T\}$;
  $\boldsymbol{z} \leftarrow \boldsymbol{z}\perp z_m$;
end
for $i = 1, 2, \ldots, T$ **do**
  for $t = 1, \ldots, \overline{T}$ **do**
    $a_i \leftarrow V(\boldsymbol{z}, \boldsymbol{r}\perp\{t\}) - V(\boldsymbol{z}, \boldsymbol{r})$;
  end
  $r_i \leftarrow \text{argmax}\{a_t | t = 1, 2, \ldots, \overline{T}\}$;
  $\boldsymbol{r} \leftarrow \boldsymbol{r}\perp r_i$;
end

Algorithm 1: Greedy Algorithm to solve $(\boldsymbol{z}, \boldsymbol{r})$ for given $\boldsymbol{y}^k$

- Any slot can generate **revisit** patients to the lab test station according to a slot-wise reentry probability $\boldsymbol{y} = (y_1, y_2, \ldots, y_T)$, with **patient type** determined by the hour of that slot. Therefore, at the lab test station, the arrival of patients from the outpatient station is exogeneously given.
- The **sojourn time** at the lab test station for revisit patients is sampled from a slot-wise distribution generated by the simulation model of the lab test station.
- (Can be removed) There is also a **walking time** associated with every revisit patient. The walking time is exogenously sampled for each slot according to the hour of that slot, and will be added to patient's sojourn time. Note that in the approximation model, the walking time will not influence the patient's arrival time at the lab test station.

And the simulaiton model for the lab test station is utilized to generate **sojourn time distribution** for revisit patients generated at each slot. We assume that:

- At the lab test station, all the patients join the same queue for routing according to their arrval time. Patients arrive at the same time will be in random order among them. Both routing and serving follow FCFS policy.

Let p denote any patient among $p_i^{j,s}$, $p_i^{j,w}$, and $p_i^{j,r}$, $j \in \{1, 2, 3, 4, 5\}, i \in \{0, 1, 2, \ldots\}$, or a virtual patient (representing an empty slot). Suppose that he receives his oupatient service at slot t and $\eta_t = 1$, indicating that he needs a revisit. Then he will arrive at the lab test station at slot $t + 1$ with patient type e, which depends on time t.

We first set a seed for each patient, which depends on his outpatient service slot t, patient type e, and outpatient doctor id d. Then we use that seed to generate a tuple of patient information from $INFO$, say $\boldsymbol{f} = (item, age, gender)$. Let function g denotes the mapping from item to its item type.

Suppose that upon his arrival, it is on weekday w, there are in total $\boldsymbol{QL} = (QL_1, QL_2, QL_3, QL_4)$ patients waiting in queue for each type at the lab test station, and there are NUM servers in service. Then the patient will be routed to a room type rt accoording to a distribution depending on $(f_2, g(f_1), t, w, \boldsymbol{QL}, NUM)$. Suppose that there are nrt number of rooms of room type rt. We further suppose that there are in total $\boldsymbol{ql} = (ql_1, ql_2, \ldots, ql_{nrt})$ patients waiting in queue for each room of room type rt, and there are num servers of room type rt in service. Then the patient will be routed to a room r accoording to a distribution depending on $(f_2, g(f_1), t, w, \boldsymbol{ql}, num)$. If the server in room r is not available, he will be added to the queue of room r.

Once the room is available, the patient will enter the service. Suppose that he enters the service at time t^{lab}. His service time is sampled from a distribution depending on (f_1, rt, t^{lab}).

Then we can calculate the total sojourn time for the patient, which is the sum of waiting time and service time. The former one depends on the patients in front of room r upon his arrival, and r depends on the patient's information, current time, weekday, number of servers, and queue length generated by patients coming previously. The later one depends on the patient's information, type of room r, and current time. Therefore, the sojourn time of the patient has nothing to do with patients coming later than him.

2.2 Performance Metrics

In the simulation model, to calculate waiting cost and completion cost, we need to record several important time stamps.

For a patient p from the outpatient station, he will be given with an attribute called "time" which records:

$$\begin{array}{l c} & \begin{array}{cccc} arrival & start & end & schedule \end{array} \\ \begin{array}{l} outpatient \\ lab \\ outpatient \\ lab \\ outpatient \end{array} & \left(\begin{array}{cccc} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & 0 \\ t_{31} & t_{32} & t_{33} & t_{34} \\ t_{41} & t_{42} & t_{43} & 0 \\ t_{51} & t_{52} & t_{53} & t_{54} \end{array} \right), \end{array} \tag{2}$$

We denote his total waiting time as $p^w_{i,j,h}$ if he is the hth patient served by outpatient doctor i on day j, then we have

$$p^w_{i,j,h} = \sum_{k=1,3,5} (t_{k2} - t_{k1})$$

, where $t_{k1} = max(t_{k-1,3}, t_{k-2,4}), k = 3, 5$. Therefore, the average waiting time W for one day is

$$W = \frac{1}{5} \frac{1}{K} \sum_{i=1}^{5} \sum_{j=1}^{K} \sum_{h} p^w_{i,j,h}$$

For each outpatient doctor d, we record his finish time for each day. We use d^f_{ij} to denote the finish time for outpatient doctor i on day j, then we have the average completion time F is

$$F = \frac{1}{5} \frac{1}{K} \sum_{i=1}^{5} \sum_{j=1}^{K} d^f_{i,j}$$

Therefore, the average total cost C is calculated as:

$$C = c_w W + c_{comp} F$$

In the approximation model,