A Markov Chain Model for Workflow Analysis in Operating Rooms

Hanyi Zheng, *Student Member, IEEE,*, Qing Wang and Jingshan Li, *Fellow, IEEE*

***Abstract*— Improving workflow efficacy in operating rooms (OR) is of significant importance for hospital management. Although extensive studies have been carried out in OR scheduling, efficient methods for workflow analysis are missing in current literature. To bridge this gap, in this paper, a Markov chain model is presented to evaluate the workflow performance in ORs. It is shown that such a model can provide efficient analysis with acceptable accuracy. In addition, a case study in a large public hospital in Beijing, China, illustrates the applicability of the method.**

I. Introduction

Operating room is one of the most cricital and expensive units in a hospital. It is estimated that about 60-70% of hospital admissions are related to surgeries, which contributes to almost 66% of total revenue and 40% of expense in a hospital [1]–[3]. Therefore, there is an increasing need to optimize surgical process and administration through effective OR management to increase utilization of medical resources, minimize patient waiting time, and reduce cost.

To improve both hospital- and patient-centric metrics, such as OR utilization, throughput, and waiting time, OR workflow management plays a key role. Although numerous studies have been carried out to simulate and optimize the performance of ORs (see, for instance, reviews [4]–[6]), most of them focus on surgery scheduling only using deterministic models. For instance, using column generation techniques, both nurse and OR scheduling processes are investigated in paper [7]. Dynamic programming and column generation algorithms are used in paper [8] to solve surgical case sequencing problem in an enumerative branch-and-price framework. Paper [9] introduces a dynamic OR scheduling method to consider specialty, surgeon, patient allocation and sequence simultaneously. To address uncertainty in OR scheduling, robust optimization and stochastic programming are used in a few studies. For example, a stochastic optimization model to develop OR schedules based on surgery duration variance is presented in paper [3]. A mixed integer linear programming model using genetic algorithm is introduced in paper [10] to study multi-period multi-resource OR schedules under uncertainty. Moreover, scheduling with uncertainties in arrivals and emergency cases with waiting time limits are discussed in papers [11] and [12], respectively. In addition, a mixed-integer quadratic model is introduced in paper [13] to optimize the tactical surgery schedule to balance the expected day-to-day occupancy of scheduled patients in the surgical intensive care unit.

However, in all these studies, the detailed settings and operation policies inside ORs are often ignored. This make it difficult to standardize the process, and propose or implement improvement strategies to increase efficiency, quality and safety [1]. Except for a few discrete-event simulation studies (such as review [14] and representative papers [15]–[20]), the current literature does not provide sufficient research results addressing the workflow issues. Therefore, there is a need to develop an analytical model of OR workflow to incorporate the complex surgical processes and critical settings (such as procedures, staff level, and scheduling policies).

In this paper, a workflow model of surgical process is presentced and a novel system-theoretic method is developed to evaluate OR performance. As appropriately defined system states and their transitions can be introduced to describe the scenarios with multiple parallel and sequential structures, limited resourses, etc., a system-theoretic approach is developed based on Markov chain analysis to efficiently calculate system performance, such as throughput, OR and staff utilization, and investigate system properties, which are the main contributions of the paper. Such a work provides a quantitative method for hospital management to plan and schedule OR activities and improve performance metrics.

The remainder of the paper is structured as follows: Section II presents the description of surgical processes in ORs and introduces a Markov chain model. The solution approach is described in III, and analysis of system properties is conducted in Section IV. A case study at a public hosptial surgery department is introduced in Section V. Finally, conclusions and future work are discussed in Section VI. All proofs are provided in the Appendix.

II. System Model

Consider a typical surgical procedure in a Chinese hospital, which can be divided into preoperative, intraoperative, and postoperative three phases, where multiple steps are included in the procedure. There are five typical steps in many hospital ORs: pre-operative registration; tripartite verification; anesthesia; surgery; and post-operative recovery.

- Step 1: *Registration and information checking*. This will be carried out by a nurse upon a patient's arrival.
- Step 2: *Tripartite verification*. When the patient is escorted into an OR, a surgical team consisting of the surgeon, anesthetist, and OR nurses will carry out tripartite verification.
- Step 3: *Anesthesia*. Each patient will receive anesthesia by an anesthetist.

H. Zheng, Q. Wang and J. Li are with the Department of Industrial Engineering, Tsinghua University, Beijing 100084, China. zhenghy22@mails.tsinghua.edu.cn, qingwang911@tsinghua.edu.cn, jsli@tsinghua.edu.cn.

- Step 4: *Surgery*. The surgery process will start.
- Step 5: *Recovery and discharge*. A surgery patient will be sent to an available bed in PACU for recovery. If none of the beds is available, the patient will recover in the OR. Afterwards, he/she will be escorted back to ICU or ward.

By ignoring the registration and information checking process, which is seldom blocked in practice, and introducing a late arrival process to describe the delays in the registration, we obtain a simplified model as shown in Figure 1.

Remark 1: In an ideal scenario, a patient should have been arrived and been ready for entering the OR once the last patient finish his/her surgery. Due to variations, some patients may arrive at the surgical suite late so that they may not finish registration in time. Such delays can be described by the late arrival process as shown in Figure 1.

Remark 2: Many hospitals in China are in shortage of anesthetists so that each anesthetist has to take responsibility of multiple ORs simultaneously and may only be able to show up in one OR during the tripartite verification and anesthesia processes. Typically an anesthesia nurse stays with the patient all the time. Thus, in this model, the anestentist is only required as a resource during the tripartite verification and anesthesia time period. Moreover, when a patient is moved to PACU, he/she will be taken care of by the anesthesia nurses in the PACU, who usually have sufficient capacity.

To analyze such processes, the following assumptions define the surgical workflow, staff, OR and PACU capacity, as well as operation policies within the processes.

(i) A surgical suite in a hospital includes M ORs and N beds in PACU.
(ii) Each OR has a dedicated surgical team.
(iii) There are R anesthetists in the suite, each is responsible for a certain number of surgical teams. The anesthetist is called based on a first-call-first-service policy.
(iv) A patient's late arrival probability is defined by p.
(v) The processes of late arrival, tripartite verification, anesthesia, surgery, and recovery follow exponential distributions with mean time parameters τ_{arr}, τ_{tri}, τ_{ane}, τ_{sur}, and τ_{rec}, respectively.

Remark 3: The exponential assumption is introduced to make the analysis tractable. It is shown in the case study that such an assumption does not lead to large deviation in performance measures. In future work, such an assumption will be lifted to consider general distribution cases.

Let T_{or} be the average time a patient stays in the OR. In addition, introduce ρ_{or}, ρ_{ane}, and ρ_{pac} as the utilization ratio of ORs, anesthetists, and PACUs, respectively. Then T_{or}, ρ_{or}, ρ_{ane}, and ρ_{pac} are the key performance indicators of a surgical suite, and are functions of all system parameters.

$$\begin{aligned} T_{or} &= f_t(M,N,R,p,\tau_{arr},\tau_{tri},\tau_{ane},\tau_{sur},\tau_{rec}), \\ \rho_{or} &= f_u(M,N,R,p,\tau_{arr},\tau_{tri},\tau_{ane},\tau_{sur},\tau_{rec}), \\ \rho_{ane} &= f_a(M,N,R,p,\tau_{arr},\tau_{tri},\tau_{ane},\tau_{sur},\tau_{rec}), \\ \rho_{pac} &= f_p(M,N,R,p,\tau_{arr},\tau_{tri},\tau_{ane},\tau_{sur},\tau_{rec}), \end{aligned} \tag{1}$$

Then, the problem can be formulated as: *Under assumptions (i)-(v), develop a method to evaluate OR performance measures and investigate system properties.*

III. Performance Evaluation

A. States

Assumptions (i)-(v) define an ergodic stochastic process. The state of the process is denoted by

$$(n_{arr_{pac}}, n_{arr_{or}}, n_{tri}, n_{ane}, n_{sur}, n_{rec_{pac}}, n_{rec_{or}}),$$

where n_{tri}, n_{ane}, and n_{sur} represent the number of patients in tripartite verification, anaesthesia (including preparation), and surgery processes, respectively, $n_{rec_{pac}}$ and $n_{rec_{or}}$ describe the number of patients who are recovering in PACU and OR, respectively, and $n_{arr_{pac}}$ and $n_{arr_{or}}$ denote the number of patients who arrive late while prior patients are recovering in PACU and OR, respectively. Then

$$n_{arr_{pac}}, n_{arr_{or}}, n_{tri}, n_{ane}, n_{sur}, n_{rec_{or}} \in \{0,1,2,\ldots,M\}, \\ n_{rec_{pac}} \in \{0,1,2,\ldots,N\}.$$

It can be shown that the total number of states is $(M+1)^6(N+1)$. Since a new patient cannot enter the OR until the prior patient leaves, there exist nonfeasible states. Thus, only the states satisfying the following condition are feasible:

$$\begin{aligned} & n_{arr_{pac}} + n_{arr_{or}} + n_{tri} + n_{ane} + n_{sur} = M, \\ & n_{arr_{or}} = n_{rec_{or}}. \end{aligned} \tag{2}$$

Then, we obtain

Proposition 1: Under assumptions (i)-(v), the number of feasible states can be calculated as

$$K = \frac{(N+1)\left[\prod_{i=1}^{4}(M+i)\right]}{24}. \tag{3}$$

Proof: See the Appendix. ■

Then the l-th state, $l \in \{1,2,\ldots,K\}$, can be denoted as

$$S^l = (n^l_{arr_{pac}}, n^l_{arr_{or}}, n^l_{tri}, n^l_{ane}, n^l_{sur}, n^l_{rec_{pac}}, n^l_{rec_{or}}), \tag{4}$$

where superscript l represents the state number.

B. Transitions

Introduce transition rate β^l_{ij} to describe the transitions between the states i and j, where i and j are the steps in the workflow. By considering the following events, we can derive the transition rates,

- First, consider the transition from tripartite verification to anesthetization. The anesthetist is called based on first-call-first-service policy, from state S^l to state S^m,

$$\begin{aligned} S^m = (& n^l_{arr_{pac}}, n^l_{arr_{or}}, n^l_{tri}-1, n^l_{ane}+1, n^l_{sur}, \\ & n^l_{rec_{pac}}, n^l_{rec_{or}}), \end{aligned}$$

the transition rate $\beta^l_{tri,ane}$ can be derived as

$$\beta^l_{tri,ane} = \begin{cases} \frac{n^l_{tri}}{\tau_{tri}}, & \text{if } n^l_{tri}+n^l_{ane} < R, \\ \frac{R}{\tau_{tri}} \cdot \frac{n^l_{tri}}{n^l_{tri}+n^l_{ane}}, & \text{if } n^l_{tri}+n^l_{ane} \geq R, \end{cases} \tag{5}$$

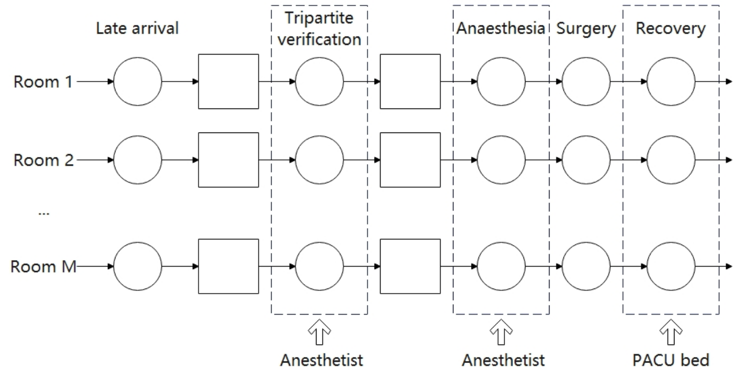


Fig. 1: Workflow in ORs

- Next, for the transition from anesthetization to surgery, state S^m is defined by

$$S^m = (n^l_{arr_{pac}}, n^l_{arr_{or}}, n^l_{tri}, n^l_{ane} - 1, n^l_{sur} + 1, n^l_{rec_{pac}}, n^l_{rec_{or}}).$$

Then the transition rate $\beta^l_{ane,sur}$ can be evaluated as

$$\beta^l_{ane,sur} = \begin{cases} \frac{n^l_{ane}}{\tau_{tri}}, & \text{if } n^l_{tri} + n^l_{ane} < R, \\ \frac{R}{\tau_{tri}} \cdot \frac{n^l_{ane}}{n^l_{tri}+n^l_{ane}}, & \text{if } n^l_{tri} + n^l_{ane} \geq R, \end{cases} \tag{6}$$

- When a patient finishes surgery and starts recovery, a bed in PACU will be used first if it is available. In addition, whenever a patient finishes surgery, the next patient will be ready to start the tripartite verification if he/she is not late. Thus, for the transition from surgery to recovery in PACU when the next patient is not late, then state S^m is

$$S^m = (n^l_{arr_{pac}}, n^l_{arr_{or}}, n^l_{tri} + 1, n^l_{ane}, n^l_{sur} - 1, n^l_{rec_{pac}} + 1, n^l_{rec_{or}}).$$

and the transition rate β^l_{sur,pac_t} will be

$$\beta^l_{sur,pac_t} = \begin{cases} \frac{n^l_{sur}(1-p)}{\tau_{sur}}, & \text{if } n^l_{rec_{pac}} < N, \\ 0, & \text{if } n^l_{rec_{pac}} \geq N, \end{cases} \tag{7}$$

- When a patient finishes surgery and starts recovery in a PACU but the next patient is late. State S^m is

$$S^m = (n^l_{arr_{pac}} + 1, n^l_{arr_{or}}, n^l_{tri}, n^l_{ane}, n^l_{sur} - 1, n^l_{rec_{pac}} + 1, n^l_{rec_{or}}).$$

The transition rate β^l_{sur,pac_a} is

$$\beta^l_{sur,pac_a} = \begin{cases} \frac{n^l_{sur}p}{\tau_{sur}}, & \text{if } n^l_{rec_{pac}} < N, \\ 0, & \text{if } n^l_{rec_{pac}} \geq N, \end{cases} \tag{8}$$

- If the patient finishes surgery and has to stay in the OR for recovery and the next patient arrives, state S^m is

$$S^m = (n^l_{arr,pac}, n^l_{arr_{or}} + 1, n^l_{tri}, n^l_{ane}, n^l_{sur} - 1, n^l_{rec_{pac}}, n^l_{rec_{or}} + 1).$$

Then, the transition rate $\beta^l_{sur,or}$ can be calculated as

$$\beta^l_{sur,or} = \begin{cases} \frac{n^l_{sur}}{\tau_{sur}}, & \text{if } n^l_{rec_{pac}} \geq N, \\ 0, & \text{if } n^l_{rec_{pac}} < N, \end{cases} \tag{9}$$

- In case of late arrivals, if the prior patient has stayed in PACU for recovery, then from late arrival to tripartite verification, state S^m is

$$S^m = (n^l_{arr_{pac}} - 1, n^l_{arr_{or}}, n^l_{tri} + 1, n^l_{ane}, n^l_{sur}, n^l_{rec_{pac}}, n^l_{rec_{or}}).$$

The transition rate $\beta^l_{pac,tri}$ is

$$\beta^l_{pac,tri} = \frac{n^l_{arr_{pac}}}{\tau_{arr}}. \tag{10}$$

- When the patient arrives late and the prior patient has stayed in OR for recovery, the scenario becomes more complicated such that the prior patient leaving the OR and the next patient arriving at the OR may not occur simultaneously. To simplify the model, we assume the late timing for both events. Thus, for the transition from late arrival to tripartite verification, state S^m is

$$S^m = (n^l_{arr_{pac}}, n^l_{arr_{or}} - 1, n^l_{tri} + 1, n^l_{ane}, n^l_{sur}, n^l_{rec_{pac}}, n^l_{rec_{or}} - 1).$$

The transition rate $\beta^l_{or,tri}$ is

$$\beta^l_{or,tri} = \min\Big(\frac{n^l_{arr_{or}}}{\tau_{arr}}, \frac{n^l_{rec_{or}}}{\tau_{rec}}\Big)p + \frac{n^l_{rec_{or}}}{\tau_{rec}}(1-p). \tag{11}$$

- Finally, considering the scenario that a recovered patient in PACU leaves the suite, state S^m is

$$S^m = (n^l_{arr_{pac}}, n^l_{arr_{or}}, n^l_{tri}, n^l_{ane}, n^l_{sur}, n^l_{rec_{pac}} - 1, n^l_{rec_{or}}),$$

Then the transition rate $\beta^l_{rec_{pac}}$ is

$$\beta^l_{rec_{pac}} = \frac{n^l_{rec_{pac}}}{\tau_{rec}}. \tag{12}$$

C. Balance Equations

To derive the expressions of performance measures, balance equations need to be developed first. Consider set

$$\Theta = \{or, tri; pac, tri; tri, ane; ane, sur; sur, pac_t; sur, pac_\alpha; sur, or; rec_{pac}\}. \quad (13)$$

Then,

$$\sum_{i\in\Theta} \beta_i^l P_l = \sum_{j\in\Theta} \beta_j^{k_j^l} P_{k_j^l}, \quad l, k_j^l \in \{1, 2, \ldots, K\}, \quad (14)$$

where superscript k_j^l is the state index that indicates after transition j the system will jump from state k_j^l to state l.

Next, compose a transition matrix Φ, where

$$\Phi(k_i^l, l) = \begin{cases} -\beta_i^{k_i^l}, & \text{if } P_{k_i^l} > 0, \\ 0, & \text{if } P_{k_i^l} = 0, \end{cases} \quad k_i^l \neq l, \quad i \in \Theta, \quad (15)$$

$$\Phi(l, l) = \sum_{i\in\Theta} \beta_i^l. \quad (16)$$

Moreover, the sum of all probabilities equals to 1, thus

$$\sum_{i=1}^{K} P_i = 1. \quad (17)$$

Using the first $K-1$ rows in equation (14) and equation (17), we obtain a new matrix Γ such that

$$\Gamma(i, j) = \Phi(i, j), \quad i = 1, \ldots, K-1, \quad j = 1, \ldots, K,$$
$$\Gamma(K, j) = 1, \quad j = 1, \ldots, K. \quad (18)$$

Let $X = [P_1, P_2, \ldots, P_K]^T$, we have the matrix format of the balance equations

$$\Gamma^T X = Y, \quad (19)$$

where $Y = [0, 0, \ldots, 1]^T$.

Solving the balance equations, the stationary distribution X can be calculated.

Proposition 2: Under assumptions (i)-(v), the unique distribution of system states can be calculated by

$$X = (\Gamma^T)^{-1} Y. \quad (20)$$

Proof: See the Appendix. ■

D. Performance Measures

Using state distribution, the performance measures can be evaluated.

Proposition 3: Under assumptions (i)-(v), the average length of stay in OR, the utilization of OR, anesthetists, and PACU can be calculated as

$$T_{or} = \frac{\sum_{l=1}^{K} P_l (n_{tri}^l + n_{ane}^l + n_{sur}^l + n_{rec_{or}}^l)}{\sum_{l=1}^{K} P_l (\beta_{pac,tri}^l + \beta_{or,tri}^l + \beta_{sur,pac_t}^l)},$$
$$\rho_{or} = \frac{\sum_{l=1}^{K} P_l (n_{tri}^l + n_{ane}^l + n_{sur}^l + n_{rec_{or}}^l)}{M},$$
$$\rho_{pacu} = \frac{\sum_{l=1}^{K} P_l n_{rec_{pac}}^l}{N}, \quad (21)$$
$$\rho_{ane} = \frac{[\sum_{l=1}^{K} P_l \beta_{ane,sur}^l](\tau_{tri} + \tau_{ane})}{R}.$$

Proof: See the Appendix. ■

E. Model Validation

To check the accuracy of assuming the same timing of a patient's leaving OR and the next patient's arrival, the proposed model is investigated by comparing with discrete-event simulations without the assumption. The simulation experiments are carried out as follows: 1,000 data sets are randomly generated, and for each experiment, a one-year period is selected for simulation time and the average performance metrics are calculated. In all experiments, the 95% confidence intervals of major performance indicators are within 5% of their corresponding values. The results are then compared with those obtained from Proposition 3. Define

$$\delta_{t,j} = \frac{|T_{or,j}^{sim} - T_{or,j}^{cal}|}{T_{or,j}^{sim}} \cdot 100\%,$$
$$\delta_{\rho_i,j} = \frac{|\rho_{i,j}^{sim} - \rho_{i,j}^{cal}|}{\rho_{i,j}^{sim}} \cdot 100\%, \quad i \in \{or, ane, pacu\}, \quad (22)$$
$$\widetilde{\delta_k} = \frac{\sum_{j=1}^{100} \delta_{k,j}}{100}, \quad k \in \{t, \rho_{or}, \rho_{ane}, \rho_{pacu}\},$$

where superscripts "sim" and "cal" denote the results obtained by simulations and calculations, respectively, and subscript "j" represents the j-th experiment. As one can see from Table I, all average errors are less than 1%. The maximal ones are also small. The rationale of the errors may come from simplifications in the model.

TABLE I: Model accuracy

$\widetilde{\delta_{t,j}}$	0.63%	$\max_j(\delta_{t,j})$	5.51%
$\widetilde{\delta_{\rho_{or},j}}$	0.43%	$\max_j(\delta_{\rho_{or},j})$	7.70%
$\widetilde{\delta_{\rho_{ane},j}}$	0.84%	$\max_j(\delta_{\rho_{ane},j})$	7.85%
$\widetilde{\delta_{\rho_{pacu},j}}$	0.88%	$\max_j(\delta_{\rho_{pacu},j})$	7.11%

Remark 4: Note that, in the simulation model, when a patient arrives late and the prior patient is staying in the OR for recovery, there exists a time difference between the prior patient leaving and the next patient entering the OR, while, as explained before, such a changeover time is ignored in the Markov chain model. This contributes to the difference between the model results.

F. Computation Efficiency

The computation efficiency of such a method depends on the size of state space. As shown in Table II, the computation time increases with respect to the number of ORs and/or beds in PACU. When there are less than 9 ORs, all performance measures can be calculated quickly.

Note that typically the number of PACU beds is much less than the number of ORs, thus the cases where $N > M$ are not considered (denoted as "-") in the table. In addition, the average computation time is obtained from a MacBook Air with Apple M1 processor and 8GB memory for 100 experiments with randomly generated parameters.

TABLE II: Computation time for smaller surgical suites (seconds)

M \ N	1	2	3	4	5	6	7	8	9
1	0.002	-	-	-	-	-	-	-	-
2	0.006	0.009	-	-	-	-	-	-	-
3	0.015	0.023	0.031	-	-	-	-	-	-
4	0.031	0.051	0.075	0.097	-	-	-	-	-
5	0.064	0.107	0.161	0.216	0.292	-	-	-	-
6	0.125	0.218	0.342	0.493	0.673	0.867	-	-	-
7	0.233	0.438	0.720	1.084	1.542	2.105	2.772	-	-
8	0.439	0.887	1.539	2.434	3.666	5.162	6.990	9.292	-
9	0.830	1.820	3.289	5.545	8.541	12.203	16.827	22.471	29.117

However, when the number of ORs is more than 9, the computation time may explode. As shown in Table III, when there are close to 15 ORs, the computation time will be longer than 1 minute (denoted as "x") if there are multiple PACUs.

Due to computation intensity, the current Markov chain model is only suitable for a small size surgical suite. However, since in large surgical suites, the ORs are often operated based on categories or groups, each group or pod may only involve a limited number of ORs. Therefore, the large suite can be decomposed into groups and analyzed separately.

IV. System Properties

To better understand system behavior, we investigate the monotonic properties of the system. Specifically, consider a surgical suite with eight ORs, three anesthetists, and four PACU beds. The parameters are provided below:

$$p = 0.2676, \quad \tau_{arr} = 0.1332, \quad \tau_{tri} = 0.0878,$$
$$\tau_{ane} = 0.2142, \quad \tau_{sur} = 1.7950, \quad \tau_{rec} = 1.056.$$

Under this setting, numerical studies are carried out to verify the monotonic properties of the system. First, monotonicities with respect to the number of resources, such as OR, anesthetists, and PACU beds, are investigated. As shown in Figure 2, monotonicity exhibits for all numbers. When the number of ORs is increased, the LOS in ORs and the utilization of ORs, PACU, and anesthetists are all increasing due to more patients being accommodated. If more anesthetists are working in the surgical suite, as expected, the utilization of anesthetists is decreased. In addition, less waiting for anesthetists will occur so that the LOS in ORs is decreased, so does the utilization of OR. This will result in more patients recovering in PACUs, thus the utilization of PACU becomes higher. Similarly, when there are more PACU beds in the department, the PACU utilization is increased, while the LOS in ORs and the utilizations of ORs and anesthetists are all reduced.

Next, we validate the monotonicity with respect to the average service times, such as tripartite verification, anesthesia, surgery, and recovery. Again as shown in Figure 3, monotonicity is observed in all cases. Both the LOS in ORs and OR utilization are increasing when the service time is longer. The PACU utilization is increasing with respect to recovery time but decreasing with respect to all other service times. The anesthetists become more busy when tripartite verification and anesthesia take more time, but less busizer when surgery and recovery times are shorter.

Finally, we investigate the impact of late arrival rate and time. It is clear that early arrival can lead to longer LOS in OR and higher utilization of resources, while opposite results are expected when the delay time is longer (see Figure 4).

V. Case Study

To illustrate the applicability of the model, we carry out a case study in an area of a surgery department in a large public hospital in Beijing, China. Such an area is mainly responsible for a few speciaties (e.g., cardiology, Liver and Gallbladder intervention, othopedics). There are five ORs, 3 PACU beds, and two anesthetists within the area. The distribution of surgery types and ratios among the five ORs are given in Table V. Note that other surgery types with very few surgeries in Rooms 3-5 are ignored.

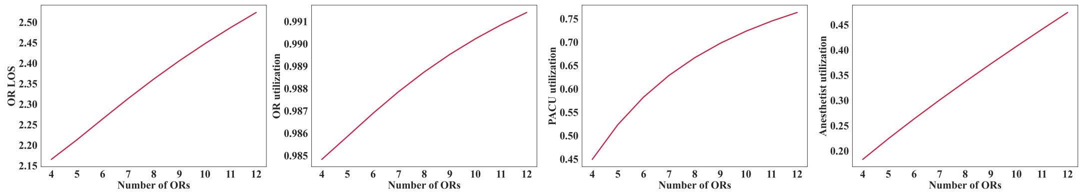
Using the Markov chain model introduced in this paper, the surgery types are generated based on their distributions in each room, and the anesthesia types are assigned for each surgery based on the probability ratio for each surgery type. As the hospital was not operated in full capacity during pandemic, while the analytical model assumes full capacity, the results obtained from the model are compared with the results regenerated using simulations to represent the same scenarios, which are developed and validated based on actual parameters calculated from the collected data in the hospital. Although the model setting is not the same, from Table V, it can be seen that the analytical model provides acceptable estimates of system performance.

VI. Conclusions and Future Work

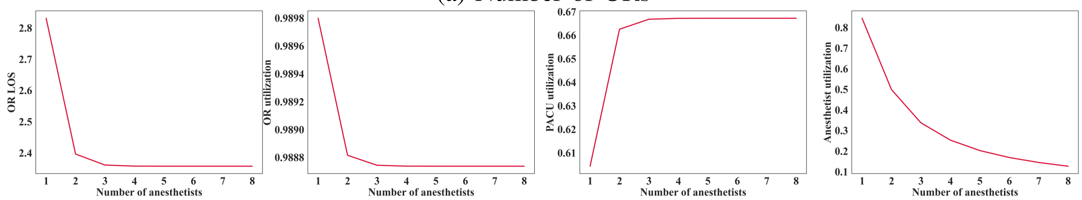
In this paper, a Markov chain model is introduced to study the workflow in ORs. Using such a model, the performance measures, such as length of stay in ORs, utilization ratios of anesthetists, OR and PACU, can be evaluated quickly and accurately. A case study at a public hospital illustrates the applicability of the model. Thus, such a study can provide quantitative predictions for planning and improvement in a hospital surgery department.

The future work can be directed to the following topics:

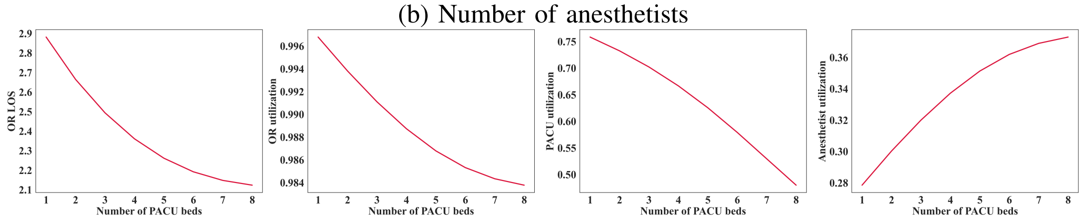
- Developing methods to study surgical suites with larger capacities, such as suites with up to 50 ORs.
- Generalizing the model to ORs with different settings.



(a) Number of ORs

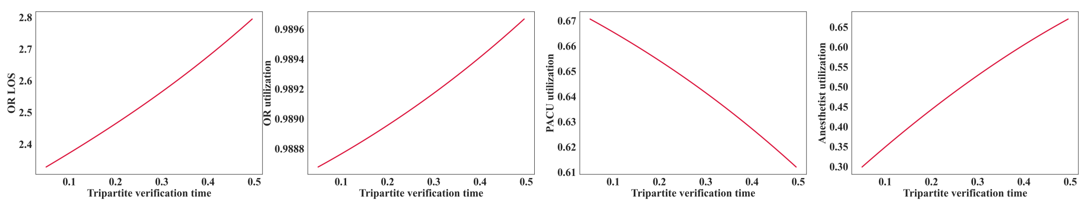


(b) Number of anesthetists

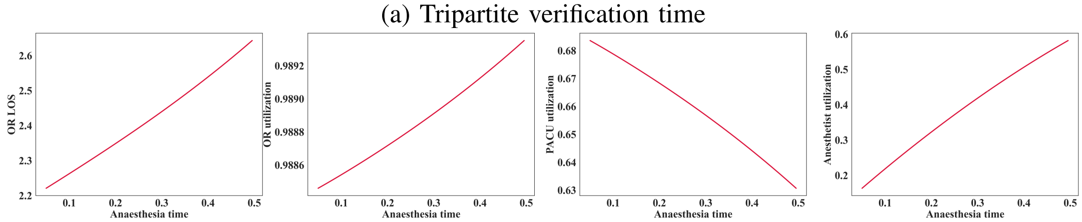


(c) Number of PACU beds

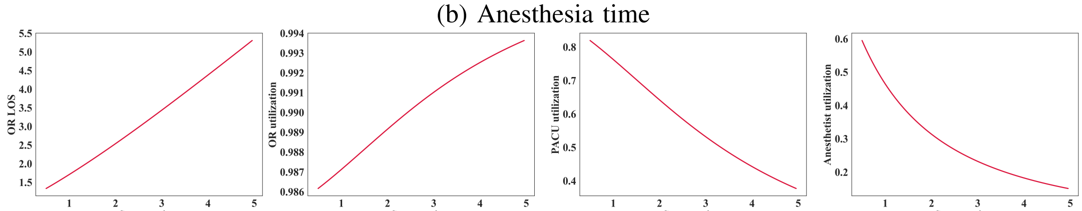
Fig. 2: Monotonicity with respect to number of resources



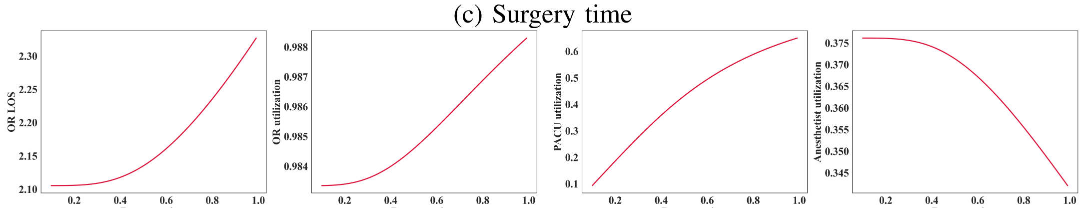
(a) Tripartite verification time



(b) Anesthesia time



(c) Surgery time

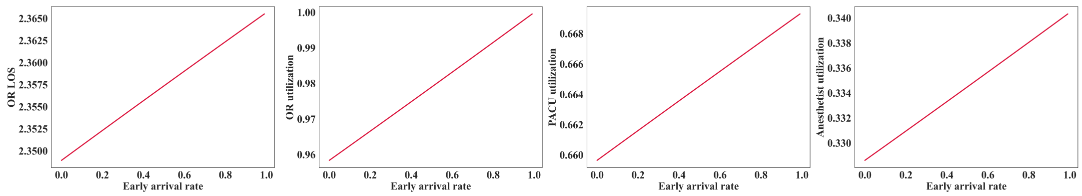


(d) Recovery time

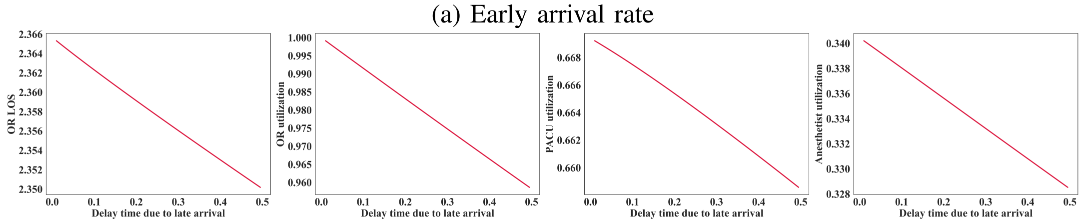
Fig. 3: Monotonicity with respect to service times

TABLE III: Computation time for larger surgical suites (seconds)

M \ N	1	2	3	4	5	6	7	8	9	10
10	1.586	3.744	7.145	12.231	18.944	27.987	39.261	53.687	x	x
11	3.031	7.578	15.169	25.994	41.773	x	x	x	x	x
12	5.786	15.391	30.671	54.962	x	x	x	x	x	x
13	10.844	29.102	x	x	x	x	x	x	x	x
14	20.062	56.803	x	x	x	x	x	x	x	x
15	35.828	x	x	x	x	x	x	x	x	x



(a) Early arrival rate



(b) Late arrival time

Fig. 4: Monotonicity with respect to late arrival rate and time

TABLE IV: ORs and specialty distribution

Room	Specialty							
	A	B	C	D	E	F	G	H
1	100%	-	-	-	-	-	-	-
2	-	100%	-	-	-	-	-	-
3	-	-	11%	7%	19%	7%	39%	17%
4	-	-	72%	-	2%	-	24%	2%
5	-	-	28%	-	23%	1%	-	48%

TABLE V: ORs and specialty distribution

	Analytical	Simulation	Difference
T_{or}	1.2412	1.27	2.27%
ρ_{or}	0.9893	0.942	5.02%
ρ_{ane}	0.2571	0.239	10.18%
ρ_{pacu}	0.3601	0.316	13.96%

- Extending the model to consider non-exponential distributions of service times.
- Applying the model in various surgery departments.

Appendix: Proofs

Proof of Proposition 1: Introduce function $f_M(m, j)$ to denote the number of feasible states for a surgical center with M ORs, when there are j non-recovery processes inside the ORs and there are m patients in the first process. Note that tripartite verification, anaesthesia, and surgery are the non-recovery processes in the ORs, so that m and j take values $m = 0, 1, 2, \dots, M$, and $j = 1, 2, 3$. Then, from (4), a feasible state S^l needs to satisfy constraint (2), so that when $m = M$ and $j = 1$, we obtain

$$n^l_{arr_{or}} = n^l_{rec_{or}} = 0, \text{ and } n^l_{arr_{pac}} + n^l_{arr_{or}} = 0,$$

and the number of feasible states depends on the number of patients in PACU, thus,

$$f^l_M(M, 1) = N + 1.$$

When there are m patients in non-recovery processes in ORs, and still $j = 1$, the constraint changes to

$$n^l_{arr_{pac}} + n^l_{arr_{or}} = M - m,$$

which presents $M - m + 1$ cases. Combining with $N + 1$ PACU cases, we obtain

$$f^l_M(m, 1) = (N + 1)(M - m + 1).$$

When $j = 2$, if one non-recovery process is fixed, the number of feasible state reduces to $f^l_M(m, 1)$. Thus, we have

$$f^l_M(m, 2) = \sum_{i \geq m} f^l_M(i, 1).$$

Similarly, when $j = 3$, it follows that

$$f^l_M(m, 3) = \sum_{i \geq m} f^l_M(i, 2).$$

Let $k = M - m$, and denote

$$f^l_M(m, 1) = a_k, \quad f^l_M(m, 2) = b_k, \quad f^l_M(m, 3) = c_k.$$

Then we obtain

$$a_i = (i+1)(N+1),$$
$$b_i = \sum_{j=0}^{i} a_j = \frac{(N+1)(i+1)(i+2)}{2},$$
$$c_i = \sum_{j=0}^{i} b_j = \frac{(N+1)(i+1)(i+2)(i+3)}{6}.$$

Finally, the total number of feasible states is

$$K = \sum_{j=0}^{M} c_j = \frac{(N+1)\left[\prod_{i=1}^{4}(M+i)\right]}{24}.$$

■

Proof of Proposition 2: Assumption (i)-(v) define an irreducible Markov chain with a finite number of states so that a unique solution exists, which can be calculated by solving balance equation (19). ■

Proof of Proposition 3: By Little's law,

$$T_{or} = \frac{WIP}{TP}, \quad \rho_{or} = \frac{WIP}{M},$$

where TP and WIP are the throughput and work-in-process (i.e., average number of patients in the surgical center), and

$$TP = \sum_{l=1}^{K} P_l(\beta^l_{pac,tri} + \beta^l_{or,tri} + \beta^l_{sur,or}),$$
$$WIP = \sum_{l=1}^{K} P_l(n^l_{tri} + n^l_{ane} + n^l_{sur} + n^l_{rec_{or}}).$$

Similarly, let WIP_{pacu} be the average number of patients recovering in PACU, WIP_{t+a} be the average number of patients in tripartite verification and anesthetization processes, TP_{t+a} be the throughput of anesthetization, we obtain

$$WIP_{pacu} = \sum_{l=1}^{K} P_l n^l_{rec_{pac}},$$
$$WIP_{t+a} = \sum_{l=1}^{K} P_l(n^l_{tri} + n^l_{ane}),$$
$$TP_{t+a} = \sum_{l=1}^{K} P_l \beta_{ane}.$$

Thus,

$$\rho_{pacu} = \frac{WIP_{pacu}}{N}, \quad T_{t+a} = \frac{WIP_{t+a}}{TP_{t+a}},$$

where T_{t+a} is the waiting time for anesthetist in tripartite verification and anesthetization processes. Since actual service time needs to exclude anesthetist waiting time, we have

$$\rho_{ane} = \frac{WIP_{t+a}}{R} \times \frac{\tau_{tri} + \tau_{ane}}{T_{t+a}}.$$

Then all performance measures can be derived. ■

References

[1] S. C. Williams, S. P. Schmaltz, D. J. Morton, R. G. Koss, and J. M. Loeb, "Quality of care in us hospitals as reflected by standardized measures, 2002–2004," *New England Journal of Medicine*, vol. 353, no. 3, pp. 255–264, 2005.

[2] B. S. Atiyeh, S. W. A. Gunn, and S. N. Hayek, "Provision of essential surgery in remote and rural areas of developed as well as low and middle income countries," *International Journal of Surgery*, vol. 8, no. 8, pp. 581–585, 2010.

[3] B. Denton, J. Viapiano, and A. Vogl, "Optimization of surgery sequencing and scheduling decisions under uncertainty," *Health Care Management Science*, vol. 10, no. 1, pp. 13–24, 2007.

[4] B. Cardoen, E. Demeulemeester, and J. Beliën, "Operating room planning and scheduling: A literature review," *European Journal of Operational Research*, vol. 201, no. 3, pp. 921–932, 2010.

[5] F. Guerriero and R. Guido, "Operational research in the management of the operating theatre: a survey," *Health Care Management Science*, vol. 14, no. 1, pp. 89–114, 2011.

[6] S. Zhu, W. Fan, S. Yang, J. Pei, and P. M. Pardalos, "Operating room planning and surgical case scheduling: a review of literature," *Journal of Combinatorial Optimization*, vol. 37, no. 3, pp. 757–805, 2019.

[7] J. Beliën and E. Demeulemeester, "A branch-and-price approach for integrating nurse and surgery scheduling," *European Journal of Operational Research*, vol. 189, no. 3, pp. 652–668, 2008.

[8] B. Cardoen, E. Demeulemeester, and J. Beliën, "Sequencing surgical cases in a day-care environment: an exact branch-and-price approach," *Computers & Operations Research*, vol. 36, no. 9, pp. 2660–2669, 2009.

[9] S. Zhu, W. Fan, T. Liu, S. Yang, and P. M. Pardalos, "Dynamic three-stage operating room scheduling considering patient waiting time and surgical overtime costs," *Journal of Combinatorial Optimization*, vol. 39, no. 1, pp. 185–215, 2020.

[10] M. M. Vali-Siar, S. Gholami, and R. Ramezanian, "Multi-period and multi-resource operating room scheduling under uncertainty: A case study," *Computers & Industrial Engineering*, vol. 126, pp. 549–568, 2018.

[11] A. Soudi, M. Heydari, and M. M. Mazdeh, "A new approach for integrated surgical procedure scheduling with arrival uncertainty," *Operational Research*, vol. 34, no. 3, p. 20, 2019.

[12] H. Miao and J.-J. Wang, "Scheduling elective and emergency surgeries at shared operating rooms with emergency uncertainty and waiting time limit," *Computers & Industrial Engineering*, vol. 160, p. 107551, 2021.

[13] S. Heider, J. Schoenfelder, S. McRae, T. Koperna, and J. O. Brunner, "Tactical scheduling of surgeries to level bed utilization in the intensive care unit," *IISE Transactions on Healthcare Systems Engineering*, vol. 10, no. 4, pp. 229–242, 2020.

[14] B. G. Sobolev, V. Sanchez, and C. Vasilakis, "Systematic review of the use of computer simulation modeling of patient flow in surgical care," *Journal of Medical Systems*, vol. 35, no. 1, pp. 1–16, 2011.

[15] B. T. Denton, A. S. Rahman, H. Nelson, and A. C. Bailey, "Simulation of a multiple operating room surgical suite," in *Proceedings of the 2006 Winter Simulation Conference*, pp. 414–424, IEEE, 2006.

[16] S. M. Ballard and M. E. Kuhl, "The use of simulation to determine maximum capacity in the surgical suite operating room," in *Proceedings of the 2006 Winter Simulation Conference*, pp. 433–438, IEEE, 2006.

[17] Q. Niu, Q. Peng, T. El Mekkawy, Y. Y. Tan, H. Bruant, and L. Bernaerdt, "Performance analysis of the operating room using simulation," *Proceedings of the Canadian Engineering Education Association*, 2007.

[18] R. A. Marjamaa, P. M. Torkki, E. J. Hirvensalo, and O. A. Kirvelä, "What is the best workflow for an operating room? a simulation study of five scenarios," *Health Care Management Science*, vol. 12, no. 2, pp. 142–146, 2009.

[19] S. Wang, V. Roshanaei, D. Aleman, and D. Urbach, "A discrete event simulation evaluation of distributed operating room scheduling," *IIE Transactions on Healthcare Systems Engineering*, vol. 6, no. 4, pp. 236–245, 2016.

[20] H. Zheng, Q. Wang, J. Shen, Y. Kong, and J. Li, "Modeling and analysis of operating room workflow in a tertiary a hospital," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7006–7013, 2022.