WILEY | Hindawi

Research Article

Dynamic Large-Scale Server Scheduling for IVF Queuing Network in Cloud Healthcare System

Yafei Li,[1] **Hongfeng Wang** [1] **Li Li,**[2] **and Yaping Fu** [3]

[1]*College of Information Science and Engineering, Northeastern University, Shenyang, China*
[2]*Xikang Healthcare Technology Co., Ltd, Shenyang, China*
[3]*School of Business, Qingdao University, Qingdao, China*

Correspondence should be addressed to Hongfeng Wang; hfwang@mail.neu.edu.cn

As one of the most effective medical technologies for the infertile patients, in vitro fertilization (IVF) has been more and more widely developed in recent years. However, prolonged waiting for IVF procedures has become a problem of great concern, since this technology is only mastered by the large general hospitals. To deal with the insufficiency of IVF service capacity, this paper studies an IVF queuing network in an integrated cloud healthcare system, where the two key medical services, that is, egg retrieval and transplantation, are assigned to accomplish in the general hospital, while the routine medical tests are assigned into the community hospital. Based on continuous-time Markov procedure, a dynamic large-scale server scheduling problem in this complicated service network is modeled with consideration of different arrival rates of multiple type of patients and different service capacities of multiple servers that can be defined as doctors of the general hospital. To solve this model, a reinforcement learning (RL) algorithm is proposed, where the reward functions are designed for four conflicting subcosts: setup cost, patient waiting cost, penalty cost for unsatisfied patient personal preferences, and medical cost of patient. The experimental results show that the optimal service rule of each server's queue obtained by the RL method is significantly superior to the traditional service rule.

1. Introduction

According to data released by the Chinese National Health Commission, the infertility rate of couples of child-bearing age in China has risen from 3% to 12%–15% in the past two decades, and infertility has become a younger development trend. It is estimated that about 50 million women are infertile, 66% of whom are below 30 years of age. For giving birth to a baby, most of them try in vitro fertilization (IVF) technology, which places very high requirements for doctors to improve success rate [1]. In China, only doctors in a limited number of general hospitals are qualified for mastering this technology. Due to the limited medical resources of general hospitals, patients have to face a very long wait for IVF procedures. However, the community hospitals are in sharp contrast with the status quo of general hospitals due to backward medical resources and low technical level of

medical staff. In very recent years, cloud healthcare has attracted considerable attention of scholars and practitioners, which can fully realize the integration of medical resources of general and community hospitals. Patients can quickly enjoy homogeneous medical services by telemedicine, also reducing the medical costs. Therefore, an integrated problem of meeting the needs of IVF patients and making the medical resources of general hospitals fully utilized is considered in this paper.

Due to the importance of improving the efficiency of system, the resource allocation or scheduling problems in a lot of complex systems, such as manufacturing system [2–4], supply chain system [5], and service system [6–8], had gained great concerns by the researchers from OR community in recent years. It is noticeable that the relevant works on medical resource allocation in healthcare system mostly focused on the scheduling of hospital beds and

operating rooms [9, 10]. Huang et al. [11] studied a patient scheduling problem in EDs, which was modeled as a multiclass queuing network with service deadlines and feedback paths. The proposed scheduling strategy could effectively alleviate the congestion of emergency departments. Erdogan et al. [12] proposed a new random integer programming model to study the dynamic sequencing and scheduling of patients. In this case, the patient's request was not in accordance with the first-come first-service (FCFS) rule, and the scheduler would reserve capacity for emergency patients. He et al. [13] applied hybrid robust random method to describe a scheduling problem of emergency room patients. Through the developed dynamic scheduling algorithm, the matching of doctors and patients was realized. Moreover, researchers have begun to focus on the influence of patient preference and selection behavior upon medical resource allocation recently [14–17]. Dogru et al. [18] developed an appointment scheduling model suitable for primary care settings and provided patients with an ideal appointment schedule based on the patient's order of appointment while meeting patient preferences. Schütz and Kolisch [19] considered a scheduling problem of allocating scarce resources to different customer categories and service types in the case of cancellation, missed appointments, and overbooking. In a patient queuing network with multiple categories, Truong [20] examined a dynamic advance scheduling problem with two patient categories and developed an optimal dynamic scheduling algorithm, which could fully adapt to daily changes in demand and capacity. An unlimited server service system with a limited-service capacity had been considered by Hassin et al. [21], in which servers were sorted by their service rate and arriving clients would join the fastest idle server.

In very recent years, regional medical cooperation has shown great advantages in improving the utilization of medical resources [22]. Especially, the cloud medical system fully realizes the sharing of medical resources among general and community hospitals. Saghafian et al. [23] studied a telemedicine system to decide whether to transfer patients to telemedicine doctors based on the knowledge of community hospital triage nurses. Rajan et al. [24] analyzed a trade-off problem between the treatment speed and quality of chronic patients in the telemedicine system and proved that the benefit-maximizing service rate gradually approaches the social optimal service rate. Erdogan et al. [12] proposed a two-stage stochastic linear model. Taking into account the patient's absentee behavior, they obtained the best planned arrival time for patients in community hospitals using the telemedicine platform and the optimal number of patients that could be arranged for daily telemedicine.

It is noticeable that most solution methods in the abovementioned works focused on the metaheuristic approaches [25–29]. The purpose of this paper is to provide an optimal dynamic resource scheduling rule and service order based on the real-time state of the system instead of the FCFS rule. For this optimal scheduling problem of dynamic tasks, many scholars had adopted reinforcement learning (RL) methods. Huang et al. [30] and Noureddine et al. [31] proposed dynamic resource allocation algorithms based on

RL method to optimize resource allocation in real time, respectively. Xiao et al. [32] developed a real-time dynamic task allocation algorithm based on Q-learning. The algorithm was not limited to adapt to its own task arrival process but also fully considered the influence of other agents on the task flow. Asghari et al. [33] proposed a RL-based resource allocation method in order to reduce the cost of system and improve the utilization of resource. Wauters et al. [34] developed a learning-based resource scheduling optimization method to minimize the average delay and total completion time of the project. This paper studies the queuing processes of egg retrieval and transplantation of infertile patients with different arrival rates, and then a dynamic resource allocation algorithm based on reinforcement learning is proposed. Considering the personal preference of patients, the algorithm is able to search for the optimal allocation plan of doctors in general hospitals and the optimal service rules of each doctor's cohort in order to minimize the average total cost of patients.

The rest of this paper is organized as follows. Section 2 presents a detailed description of the IVF process under the telemedicine system. In Section 3, the queuing problem of the IVF in the general hospital is formally defined along with notation, Markov model, objectives, and some preliminary analysis. A resource scheduling algorithm based on reinforcement learning is proposed in Section 4, and the experimental study is presented in Section 5. Finally, concluding remarks are presented in Section 6, followed by some directions for future work.

2. IVF Process in Cloud Healthcare System

The procedure of IVF is a complex and multistage process. In order to improve the utilization rate of resources and service quality, cooperation between general hospitals and community hospitals is necessary. The preexamination process of IVF can be carried out in community hospitals because of its easily learned techniques. Then, through the stimulation of drugs to promote women's ovulation process, once the egg is mature, the doctor will schedule the surgery for her to obtain ovum. After the fertilized egg develops healthily into an embryo, the cultivation stage for embryo is carried out in the community hospital, where the doctors in general hospitals use the telemedicine platform to interpret the embryo report and watch the results of fertilization, division, and blastocyst culture for community hospital doctors. After the embryo sac matures, the patient will return to the general hospital for embryo transfer. Then, they returned to the community hospital for pregnancy examination and they do not need to return to the general hospital because of the remote consultation with doctors in the general hospital through the telemedicine platform. Telemedicine can effectively realize the hierarchical medical services and also can be a great approach to save medical costs and reduce waiting time of IVF patients. Patients in community hospitals can enjoy the excellent services of general hospital just through telemedicine system. We give the IVF flow chart, depending on the actual telemedicine IVF process in the cloud healthcare system (see Figure 1).
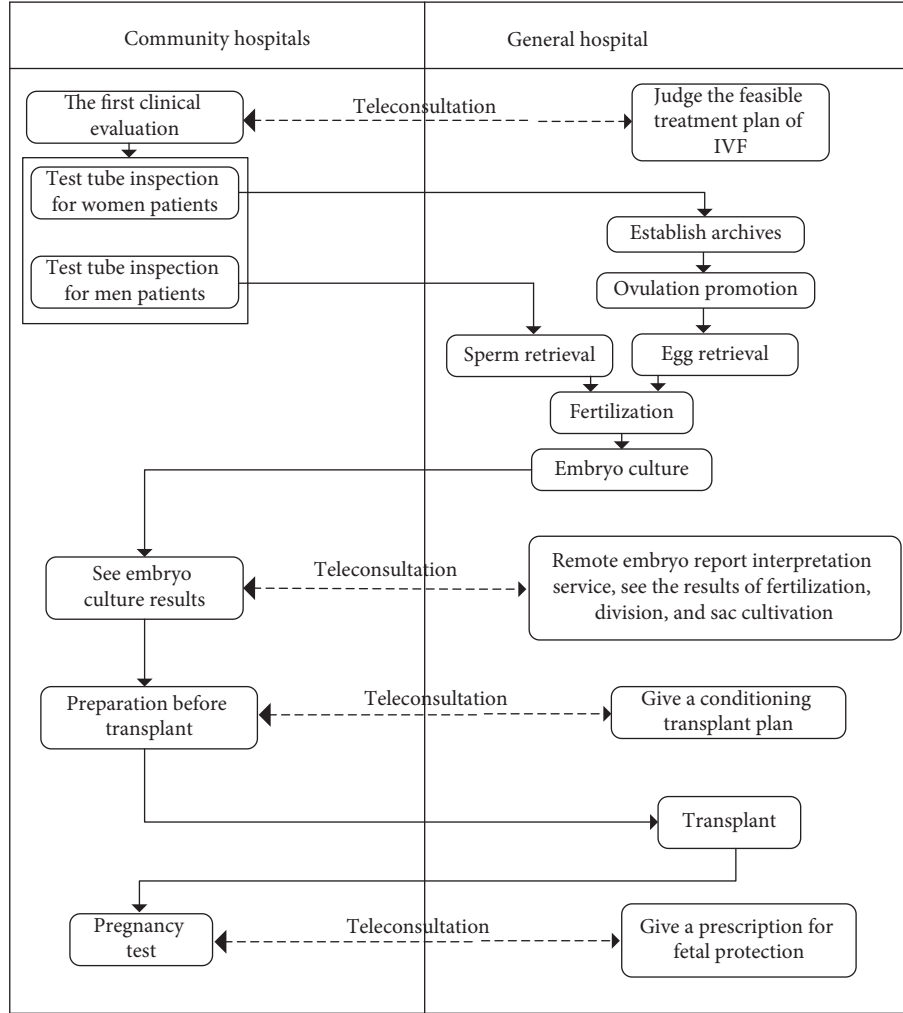
FIGURE 1: The IVF flowchart in the telemedicine system.

As one of the top-quality scarce resources, doctors in general hospitals have very close relationship with the health present situation. It is obvious from Figure 1 that doctors in general hospitals are the bottleneck that severely restricts the cloud healthcare system for operation, and they only play important roles in the egg retrieval process and the transplant process. Therefore, we just take these two queuing processes into account. The feature of the IVF procedure is time-consuming, so, in the proposed short-term resource scheduling problem, egg retrieval and transplantation are regarded as two different types of patients, and each has its own update arrival process as well as the general service time distribution. After completing the corresponding IVF process in community hospitals, patients line up to enter the queue of general hospitals for egg retrieval or transplant surgery.

3. Dynamic Resource Scheduling Model of IVF

The main problems we tackle in this paper are how to allocate the two types of infertility patients to the service queues of different doctors according to the patients' choice preferences and how to determine the better service rules of each queue. The current optimization and scheduling practice has some obvious shortcomings. First, the service order is based on the rule of FCFS. Second, when considering multiple servers, patients' choice preferences are often neglected or taken into account exactly. Therefore, we formulate a Markov model to improve the current scheduling process addressing each shortcoming. We can fill the gap in the exciting literature on scheduling problem of IVF queuing network.

In this section, we simplify the queuing process of infertile patients in general hospitals to the queuing network shown in Figure 2 according to the IVF process of the cloud medical system in Figure 1. In the cloud medical system, egg retrieval patients and transplant patients may be from different community hospitals, and they share the resource pool of general hospitals.

The notations are given in Table 1.

We can get that the arrival rate of type 1 patients and type 2 patients in general hospital are, respectively, $\lambda_1(t) = \sum_{h=1}^{H} \lambda_{h,1}(t)$ and $\lambda_2(t) = \sum_{h=1}^{H} \lambda_{h,2}(t)$. The service rates of the system serving type 1 and type 2 patients are
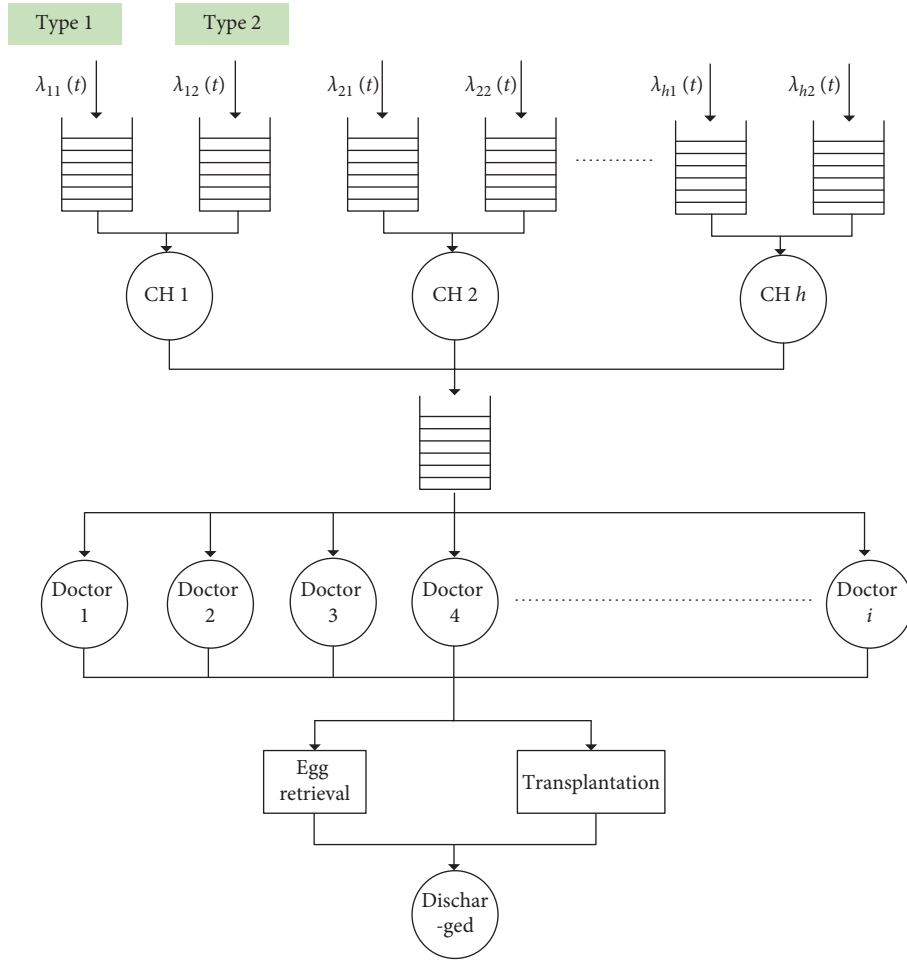
FIGURE 2: The queuing network of infertile patients in general hospital.

TABLE 1: Notation.

Sets:	
I	A set of specialists in general hospital
J	A set of types of infertility patients
H	A set of community hospitals
T	A set of time periods in the planning cycle
Indices:	
i	Index of specialists in general hospital
j	Type index of infertile patients
h	Index of community hospitals
t	Index of time periods in the planning cycle
$d(i)$	Index of patients served by doctor i
Parameters:	
$\mathscr{E}_1(t)$	The waiting time cost of the patients in the general hospital during period t
$\mathscr{E}_2$	The setup cost between disease categories 1 and 2 for specialists
$\mathscr{E}_3$	Penalty costs of unmet infertility patients' personal preferences
c_i	Unit medical cost of doctor i
π	Service rules for each doctor queue
$\lambda_{h,j}(t)$	The arrival rate of type j patient in community hospital h at the beginning of the tth time period
$\mu_{i,j}$	The service rate of type i doctor serving type j patient in the general hospital
Random variables:	
$\mathscr{J}_i(t)$	Number of service settings of type i doctor in the tth time period
Decision variables:	
$z_{d(i)}(t)$	In the tth time period, 0 if the personal preference of patient $d(i)$ served by doctor i is met and 1 otherwise
$\mathscr{Y}_{d(i)}(t)$	In the tth time period, 1 if doctor i is assigned to patient $d(i)$ and 0 otherwise

$\mu_1 = \sum_{i=1}^{I} \mu_{i,1}$ and $\mu_2 = \sum_{i=1}^{I} \mu_{i,2}$, respectively. The arrival rate and the service rate are given based on a real data from the actual cloud healthcare system. We model the queuing process of infertile patients in general hospital as the Markov decision process. We presented the state transition diagrams of the two types of patients at different time periods as Figure 3.

In Figure 3, the probability that the number of transplanted patients varies from $n+1$ to n ($0 \leq n \leq N$) is $\sum_{i=1}^{I} y_{i2}\mu_2$ in the fixed number of egg retrieval patients, because, in different planned time periods, the service queue of each doctor may contain two types of patients or only one type of patients, or even the queue is empty. In this case, the service rate of the system is changed at any moment depending on the number of doctors serving transplant patients. At different moments, the value of $\sum_{i=1}^{I} y_{i2}\mu_2$ is different. For instance, in the tth time period,

the general hospital has a total of I available doctors resources, but only a cohort of 4 doctors have 2 types of patients. Then, given the number of patients in one type, the probability that the total number of patients varies from $n+1$ to n is $4\mu_2$. In the $t+1^{\text{th}}$ time period, suppose that there are 2 types of patients in the service queue of 3 doctors; then, given that the number of patients in one type is m, the probability that the number of patients in type 2 varies from $n+1$ to n is $3\mu_2$.

In the steady state, we give the following balance equation:

For $m=0$ and $n=0$,

$$(\lambda_1(t) + \lambda_2(t))P(0,0) = \mu_1 P(1,0) + \mu_2 P(0,1). \qquad (1)$$

For $m=0$ and $n=1,2,\ldots,N-1$,

$$\left(\lambda_1(t) + \lambda_2(t) + \sum_{i=1}^{I} y_{i2}\mu_2\right)P(0,n) = \lambda_2(t)P(1,n-1) + \sum_{i=1}^{I} y_{i1}\mu_1 P(1,n)$$
$$+ \sum_{i=1}^{I} y_{i2}\mu_2 P(0,n+1). \qquad (2)$$

For $m=0$ and $n=N$,

$$\left(\lambda_1(t) + \sum_{i=1}^{I} y_{i2}\mu_2\right)P(0,N) = \lambda_2(t)P(0,N-1) + \sum_{i=1}^{I} y_{i1}\mu_1 P(1,N). \qquad (3)$$

For $m=1,2,\ldots,M-1$ and $n=0$,

$$\left(\lambda_1(t) + \lambda_2(t) + \sum_{i=1}^{I} y_{i1}\mu_1\right)P(m,0) = \lambda_1(t)P(m-1,0) + \sum_{i=1}^{I} y_{i1}\mu_1 P(m+1,0)$$
$$+ \sum_{i=1}^{I} y_{i2}\mu_2 P(m,1). \qquad (4)$$
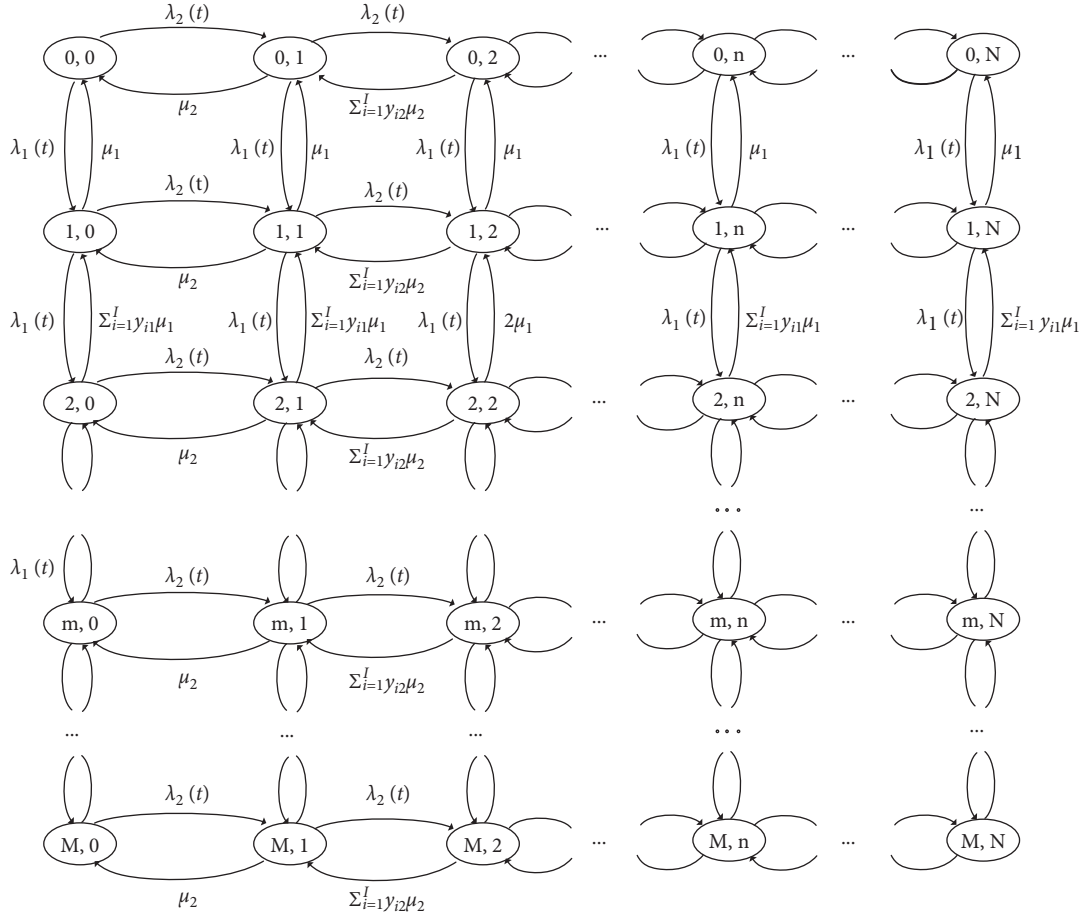
For $m=1,2,\ldots,M-1$ and $n=1,2,\ldots,N-1$,

FIGURE 3: The state transition diagrams of the two types of patients.

$$\left(\lambda_1(t) + \lambda_2(t) + \sum_{i=1}^{I} y_{i1}\mu_1 + \sum_{i=1}^{I} y_{i2}\mu_2\right)P(m,n) = \lambda_1(t)P(m-1,n) + \lambda_2(t)P(m,n-1) + \sum_{i=1}^{I} y_{i1}\mu_1 P(m+1,n)$$

$$+ \sum_{i=1}^{I} y_{i2}\mu_2 P(m,n+1).$$

(5)

For $m = 1, 2, \ldots, M-1$ and $n = N$,

$$(\lambda_1(t) + \sum_{i=1}^{I} y_{i1}\mu_1 + \sum_{i=1}^{I} y_{i2}\mu_2)P(m,N) = \lambda_1(t)P(m-1,N) + \lambda_2(t)P(m,N-1) + \sum_{i=1}^{I} y_{i1}\mu_1 P(m+1,N).$$

(6)

For $m = M$ and $n = 0$,

$$(\lambda_2(t) + \sum_{i=1}^{I} y_{i1}\mu_1)P(M,0) = \lambda_1(t)P(M-1,0) + \sum_{i=1}^{I} y_{i2}\mu_2 P(M,1).$$

(7)

For $m = M$ and $n = 1, 2, \ldots, N - 1$,

$$\left(\lambda_2(t) + \sum_{i=1}^{I} y_{i1}\mu_1 + \sum_{i=1}^{I} y_{i2}\mu_2\right) P(M, n) = \lambda_1(t)P(M - 1, n) + \lambda_2(t)P(M, n - 1) + \sum_{i=1}^{I} y_{i2}\mu_2 P(M, n + 1). \tag{8}$$

For $m = M$ and $n = N$,

$$\left(\sum_{i=1}^{I} y_{i1}\mu_1 + \sum_{i=1}^{I} y_{i2}\mu_2\right) P(M, N) = \lambda_1(t)P(M - 1, N) + \lambda_2(t)P(M, N - 1). \tag{9}$$

The infertile patients in the cloud healthcare system come from different community hospitals; in that sense, the scale of this problem is a large one. In order to effectively solve the proposed problem, the state transition diagram in Figure 3 is simplified. Given the number of patients in type 2, the state transition diagram of patients in type 1 in the tth time period is shown in Figure 4.

We assume that the upper limit of the number of patients in type 1 (i.e., egg retrieval patients) is M. Let the steady-state probability of the number (of m) of egg retrieval patients in the system be $P(Q_j = m)$. The state balance equation for the queuing process of egg retrieval patients is shown below.

For $m = 0$,

$$\lambda_1(t)P(Q_1 = 0) = \sum_{i=1}^{I} y_{i1}\mu_1 P(Q_1 = 1). \tag{10}$$

For $m = 1, 2, \ldots, M - 1$,

$$\left(\lambda_1(t) + \sum_{i=1}^{I} y_{i1}\mu_1\right) P(Q_1 = m) = \lambda_1(t)P(Q_1 = m - 1) + \sum_{i=1}^{I} y_{i1}\mu_1 P(Q_1 = m + 1). \tag{11}$$

For $m = M$,

$$\sum_{i=1}^{I} y_{i1}\mu_1 P(Q_1 = M) = \lambda_1(t)P(Q_1 = M - 1). \tag{12}$$

Define the service intensity for egg retrieval patients in the system as $\rho_1 = (\lambda_1(t)/\sum_{i=1}^{I} y_{i1}\mu_1)$; when $m = 0$, we can have

$$P(Q_1 = 0) = \frac{1}{\sum_{m=0}^{\left(\sum_{i=1}^{I} y_{i1} - 1\right)} (1/m!)(\lambda_1(t)/\mu_1)^m + \left(1/\left(\sum_{i=1}^{I} y_{i1}\right)!(1 - \rho_1)\right)(\lambda_1(t)/\mu_1)^{\left(\sum_{i=1}^{I} y_{i1}\right)}}. \tag{13}$$
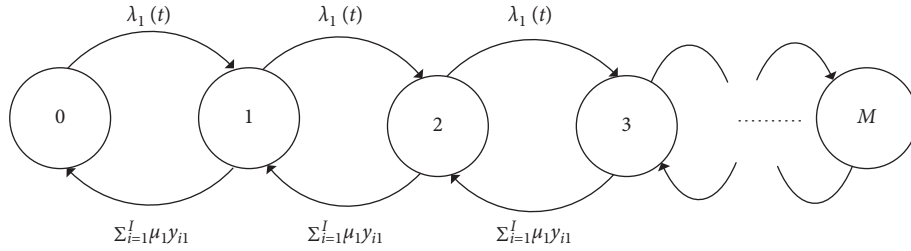
When $m \geq 1$,

FIGURE 4: The state transition diagram for type 1 patients.

$$P(Q_1 = m) = \begin{cases} \dfrac{1}{\left(\sum_{i=1}^{I} y_{i1}\right)!} \left(\dfrac{\lambda_1(t)}{\mu_1}\right)^m P(Q_1 = 0), & n \le \sum_{i=1}^{I} y_{i1}, \\[4mm] \dfrac{1}{\left(\sum_{i=1}^{I} y_{i1}\right)^{\left(m - \sum_{i=1}^{I} y_{i1}\right)} \left(\sum_{i=1}^{I} y_{i1}\right)!} \left(\dfrac{\lambda_1(t)}{\mu_1}\right)^m P(Q_1 = 0), & n > \sum_{i=1}^{I} y_{i1}. \end{cases} \tag{14}$$

In the steady state, from the state balance equation of the egg retrieval patients, the state transition probability matrix of the egg retrieval patients can be obtained:

$$\begin{bmatrix} 1 - \lambda_1(t) & \lambda_1(t) & \cdots & & 0 \\ \sum_{i=1}^{I} y_{i1}\mu_1 & 1 - \left(\sum_{i=1}^{I} y_{i1}\mu_1 + \lambda_1(t)\right) & \lambda_1(t) & \cdots & 0 \\ 0 & \sum_{i=1}^{I} y_{i1}\mu_1 & 1 - \left(\sum_{i=1}^{I} y_{i1}\mu_1 + \lambda_1(t)\right) & \lambda_1(t) & \cdots & 0 \\ \vdots & \cdots & & \ddots & \vdots \\ 0 & \cdots & & \sum_{i=1}^{I} y_{i1}\mu_1 & 1 - \sum_{i=1}^{I} y_{i1}\mu_1 \end{bmatrix}. \tag{15}$$

According to Little's formula, the queue length of egg retrieval patients in the tth time period is

$$L_1(t) = \frac{(\lambda_1(t)/\mu_1)^{\sum_{i=1}^{I} y_{i1}} \rho_1 P(Q_1 = 0)}{\left(\sum_{i=1}^{I} y_{i1}\right)! (1 - \rho_1)^2}. \tag{16}$$

The expected number of the egg retrieval patients in the tth time period is

$$L_s(t) = \sum_{m=1}^{M} m P(Q_1 = m). \tag{17}$$

The total service intensity of the system is denoted as $\rho = (\lambda_1(t) + \lambda_2(t)/\sum_{i=1}^{I} \sum_{j=1}^{J} y_{ij}\mu_j)$. The average waiting time for infertile patients in the tth time period is

$$W_1(t) = \frac{L_1(t)}{\lambda_1(t)} = \frac{(\lambda_1(t)/\mu_1)^{\sum_{i=1}^{I} y_{i1}} \rho \rho_1(Q_1 = 0)}{\left(\sum_{i=1}^{I} y_{i1}\right)! (1 - \rho_1)^2 \lambda_1(t)}. \tag{18}$$

The number of patients served in the tth time period is

$$\sum_{i=1}^{I} D_{i1}(t) = \lambda_1(t) * t - L_s(t). \tag{19}$$

The average number of customers of infertile patients in the tth time period is

$$L_1'(t) = W_1(t)\lambda_1(t) = \frac{\lambda_1(t)}{\mu_1} + \frac{(\lambda_1(t)/\mu_1)^{\sum_{i=1}^{I} y_{i1}} \rho_1 P(Q_1 = 0)}{\left(\sum_{i=1}^{I} y_{i1}\right)! (1 - \rho_1)^2}. \tag{20}$$

Furthermore, assume that the upper limit of the number of patients in type 2 (i.e., transplant patients) is N. Let the steady-state probability of the number (of n) of transplant patients in the system be $P(Q_j = n)$. The simplified state transition diagram of transplant patients is shown in Figure 5.

In the queuing system for infertile patients, the service intensity of transplant patients is defined as $\rho_2 = (\lambda_2(t) / \sum_{i=1}^{I} y_{i2}\mu_2)$. Given the number of egg retrieval patients, we can get the number of transplant patients in the tth time period. Hence, we can obtain

$$P(Q_2 = n | Q_1 = m)$$

$$= \begin{cases} \dfrac{1}{\left(\sum_{i=1}^{I} y_{i2}\right)!} \left(\dfrac{\lambda_2(t)}{\mu_2}\right)^m P(Q_2 = 0 | Q_1 = m), & m = 0, 1, \ldots, \sum_{i=1}^{I} y_{i1}, \\[4ex] \dfrac{1}{\left(\sum_{i=1}^{I} y_{i2}\right)^{\left(m - \sum_{i=1}^{I} y_{i2}\right)} \left(\sum_{i=1}^{I} y_{i2}\right)!} \left(\dfrac{\lambda_2(t)}{\mu_2}\right)^n P(Q_2 = 0 | Q_1 = m), & m = \sum_{i=1}^{I} y_{i1}, \ldots, M, \end{cases} \tag{21}$$

$$P(Q_2 = 0 | Q_1 = m)$$

$$= \begin{cases} \dfrac{1}{\sum_{n=0}^{\left(\sum_{i=1}^{I} y_{i2} - 1\right)} (1/n!) \left(\lambda_2(t)/\mu_2\right)^m + \left(1/\left(\sum_{i=1}^{I} y_{i2} - 1\right)!(1 - \rho_2)\right) \left(\lambda_2(t)/\mu_2\right)^{\left(\sum_{i=1}^{I} y_{i2}\right)}}. \end{cases}$$

From the conditional probability formula, we can deduce the steady-state probability of the number of transplant patients in the system as

$$P(Q_2 = n) = \sum_{m=0}^{M} P(Q_2 = n | Q_1 = m) P(Q_1 = m), \quad \text{for } n = 0, 1, \ldots, N. \tag{22}$$

When the number of egg retrieval patients is given, the state transition probability matrix of transplant patients can be obtained as

$$\begin{bmatrix} 1 - \lambda_2(t) & \lambda_2(t) & \cdots & 0 \\ \sum_{i=1}^{I} y_{i2}\mu_2 & 1 - \left(\sum_{i=1}^{I} y_{i2}\mu_2 + \lambda_2(t)\right) & \lambda_2(t) & \cdots & 0 \\ 0 & \sum_{i=1}^{I} y_{i2}\mu_2 & 1 - \left(\sum_{i=1}^{I} y_{i2}\mu_2 + \lambda_2(t)\right) & \lambda_2(t) & \cdots & 0 \\ & \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & & \sum_{i=1}^{I} y_{i2}\mu_2 & 1 - \sum_{i=1}^{I} y_{i2}\mu_2 \end{bmatrix}. \tag{23}$$
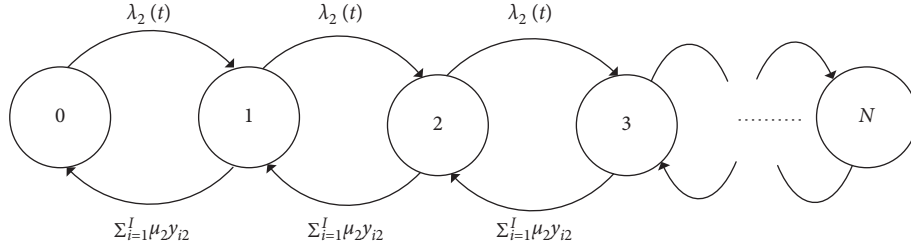
FIGURE 5: The state transition diagram for type 2 patients.

The queue length of transplant patients in the tth time period is

$$L_2(t) = \frac{(\lambda_2(t)/\mu_2)^{\sum_{i=1}^{I} y_{i2}} \rho_2 P(Q_2 = 0|Q_1 = m)}{\left(\sum_{i=1}^{I} y_{i2}\right)! (1 - \rho_2)^2}. \quad (24)$$

The expected number of transplant patients in the tth time period is

$$L_s'(t) = \sum_{n=1}^{N} n P(Q_2 = n|Q_1 = m). \quad (25)$$

The average waiting time for transplant patients in the tth time period is

$$W_2(t) = \frac{L_2(t)}{\lambda_2(t)} = \frac{(\lambda_2(t)/\mu_2)^{\sum_{i=1}^{I} y_{i2}} \rho_2 P(Q_2 = 0|Q_1 = m)}{\left(\sum_{i=1}^{I} y_{i2}\right)! (1 - \rho_2)^2 \lambda_2(t)}. \quad (26)$$

The number of transplant patients served in the tth time period is

$$\sum_{i=1}^{I} D_{i2}(t) = \lambda_2(t) * t - L_s'(t). \quad (27)$$

The average number of transplant patients in the tth time period is

$$L_2'(t) = W_2(t)\lambda_2(t) = \frac{\lambda_2(t)}{\mu_2} + \frac{(\lambda_2(t)/\mu_2)^{\sum_{i=1}^{I} y_{i2}} \rho_2 P(Q_2 = 0|Q_1 = m)}{\left(\sum_{i=1}^{I} y_{i2}\right)! (1 - \rho_2)^2}. \quad (28)$$

We consider the setup cost when the doctor switches from the current type of patients to another; the aim is to gain the largest equilibrium benefits between community hospitals and general hospital. Define the setting times of each doctor in t time period as $\mathcal{J}_i(t) = \sum_{d(i)}^{D(i)} x_{d,d-1}^i$. Let $\mathcal{D}_i(t) = \sum_{j=1}^{J} \mathcal{D}_{ij}(t)$ be the total number of two types of patients served by doctor i in the time period t. The average waiting time and queue length of the two types of infertility

patients in the tth time period are, respectively, $W(t) = W_1(t) + W_2(t)$ and $L(t) = L_1(t) + L_2(t)$. In a planning period T, the matching problem between doctors and patients can be well solved according to the model established by us. The objective function is established as follows:

$$\mathcal{F}^{\pi}(t) = \min_{t \in T} \left(\frac{1}{t} \left[\sum_{t=1}^{T} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{d(i)}^{D(i)} \left(\frac{1}{\mu_{i,j}} c_i \mathcal{D}_i(t) \mathcal{Y}_{d(i)}(t) + z_d^i(t) \mathcal{E}_3 \right) + \sum_{t=1}^{T} \mathcal{E}_1(t) W(t) L(t) + \sum_{t=1}^{T} \sum_{i=1}^{I} \mathcal{J}_i(t) \mathcal{E}_2 \right] \right). \quad (29)$$

The objective function minimizes the total costs, the first term of the objective function represents the medical cost of infertile patients, the second term is the penalty cost of unmet infertility patients' personal preferences, the third term represents the waiting time cost of egg retrieval patients and transplant patients in the system, and the last term is doctors' setup cost. In general, the above objective reflects the interests of both general hospitals and infertile patients

in the cloud medical system. In order to avoid a patient's too long waiting time in the system, a four-cost reward function based on reinforcement learning is designed.

4. Q-Learning-Based Solution Method

Existing scheduling rules seldom solve the problem of doctor resource scheduling with different service rates in cloud

medical systems. The doctor scheduling that frequently appeared in the existing related literature is based on some given sequence of the patients. In a dynamic environment, doctors' service rate and patients' waiting time are considered as important indices of medical system, especially in cloud medical system with multiple hospitals cooperating with each other. The operating efficiency and the operating costs of the scheduling results greatly varied with the different scheduling rules. Therefore, our purpose is to design an optimal scheduling rule and patient service order for each doctor queue by using reinforcement learning approach. In this section, three-stage dynamic scheduling problems are proposed. At the first step, we divide the planning cycle into hourly dynamic scheduling problems; by different arrival rate and different service rate, we need to decide which patients are assigned to one of the doctors. At the second step, the service rule is presented based on the learning strategies we design. At the third step, we need to make decisions about the number of doctors serving different types of patients from community hospitals in different time periods.

Reinforcement learning is an online actor critic method in machine learning (Sutton et al. [35] and Gao et al. [36]), which obtains certain rewards through interaction with the environment and ultimately maximizes long-term returns. The typical learning algorithm is to update the current state-action pair based on the observed reward and the next state-action pair. Combined with the research questions in this article, we give the algorithm framework based on reinforcement learning in Algorithm 1.

In the IVF queuing system based on the cloud medical system studied in this paper, we define that the state space of the system is composed of the number of patients of the two types at different times and the busyness of the respective queues of doctors in general hospitals. Egg retrieval and transplant patients can only make appointments for related operations within the given appointment time period. We give the system the number of appointments allowed on the day and the strict upper limit of doctor resources. Infertility is a special disease, so the operation must be completed in a given time, even if overtime doctors also have to complete all operations. The average cost of IVF operation much surpasses common operation. Therefore, choosing a method is a key step in cost-saving and drives higher operational efficiency.

Reinforcement learning is an effective dynamic programming method to solve dynamic scheduling problems, and its basic idea is shown in Algorithm 1. It can continuously train based on data to obtain accurate responses to the environment. The main core of reinforcement learning is the design of the reward functions which are guided by the learning system targeted at minimizing the average total cost of the system. However, our total cost consists of four parts: the average service cost, the average waiting time cost, the average setup cost, and the average penalty cost for unmet infertility patients. In order to maximize the long-term total revenue, according to the greedy strategy, we set the reward functions in different states for the four subcosts. After each step is executed, the

cumulative reward score of each subcost is treated as the total reward score obtained after the current action is executed, which ultimately maximizes the cumulative reward. The proposed reward functions balance the interests of both doctors and patients. In different planning time periods, different reward and punishment strategies are set according to the state of the system. According to the Markov decision model established in the previous section, the possible state of the system at each moment can be obtained, and, by using the trial-and-error method, we explore all the possible behaviors generated by the current state to find the current maximum return: $\max_{a'} Q(s', a')$.

5. Experimental Study

In this section, we discuss our findings. In particular, we describe the effectiveness of the proposed learning algorithm for reducing the waiting time and the medical cost. The scheduling rules and matching results of patients and doctors can simplify the complex scheduling problem. The IVF queuing network we studied includes one general hospital and 30 community hospitals. The problem raised is to serve hundreds of patients of two different types on six servers (general hospital doctors) with different service rates. As discussed earlier, the patients demand scenarios derived from actual demand for the sake of ensuring the veracity and reliability of this experiment.

This paper studies a resource scheduling problem within a day, where the allowable appointment time period is from 8:00 to 12:00. We divide the one-day scheduling problem into hourly subscheduling problems. We apply the sequence-based scheduling method and the learning algorithm to compare the efficiency and medical costs. Figure 6 shows a scheduling Gantt chart based on appointment orders (FCFS) of two types of patients in six parallel service desks. If the service desk is free, the patients will automatically join the queue. We are interested in how much better the overall results can be when patients' choice preference is not completely considered. In this experiment, idle servers provide services to patients in the order of appointment. However, these patients with personal preferences have to wait until the doctors of their choice are available. Finally, we calculate that the objective function value under this rule is 50816 yuan.

In Figure 7, we conduct 50 rounds of learning based on the Q-learning algorithm, where the abscissa represents the number of learning rounds and the ordinate represents the total reward value corresponding to each round. Obviously, the learning result curve of the first 40 times is very volatile, because Q-learning algorithm tries to find a better result to balance the four subgoals. The curve finally converges to near 45500, which shows that the results of Q-learning algorithm are better than those of FCFS. Figure 8 is the scheduling Gantt chart given by the 50th round of learning results, which not only shows the doctors and patients match relations but also reflects the patients' personal choice behavior. In order to reduce the long waiting time of patients, the service sequence of patients who arrive at the current time is arranged downstream of the patients who have made

Letter explanation: $Q(s, a)$: learning table about s and a;
   s: the state after each step is executed
   a: predict the action to be performed next
Require: initialize $Q(s, a)$ arbitrary value;
  Repeat (for each episode):
    Initialize s
    Repeat (for each step of episode):
      $A \longleftarrow$ action given by strategy π for s in the learning table $Q(s, a)$.
      Take action A, observe the next state s' and reward r
      $Q(s, a) \longleftarrow Q(s, a) + \alpha[r + \gamma^* \max_{a'} Q(s', a') - Q(s, a)]$
      $s \longleftarrow s'; a \longleftarrow a';$
    Until s is terminal

ALGORITHM 1: Algorithm framework based on reinforcement learning.

    — Egg retrieval patient
    — Transplant patient
    — Egg retrieval patient with personal preferences
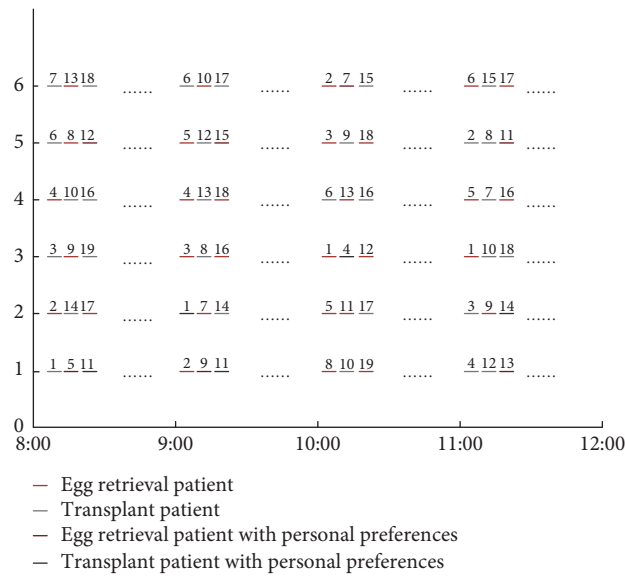   — Transplant patient with personal preferences

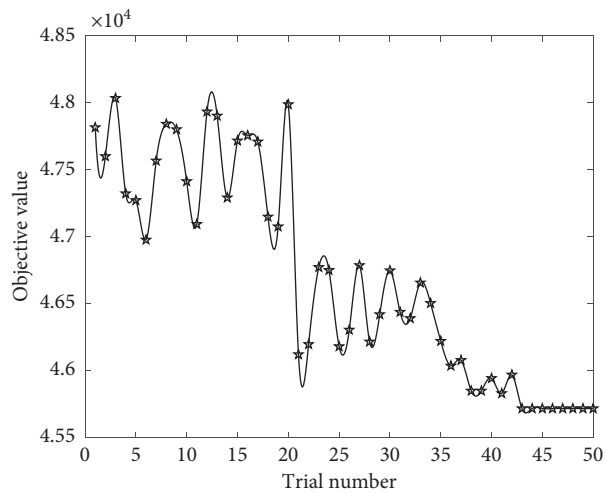FIGURE 6: Scheduling Gantt diagram based on sequence rules.

FIGURE 7: Performance curve of Q-learning scheduling approach.
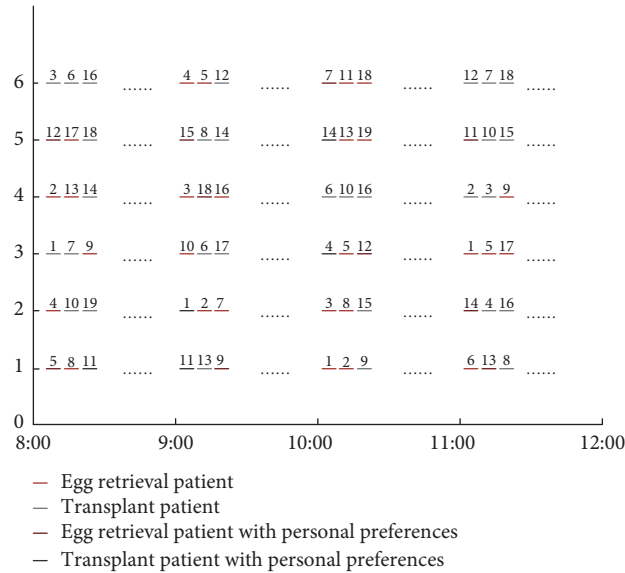
FIGURE 8: Scheduling Gantt diagram based on Q-learning.

an appointment in the previous time period. For example, for patients who make an appointment at 9 : 00–10 : 00, the Q-learning algorithm determines their service order and which doctor they should be served by, but their service order is scheduled after the patients who make an appointment at 8 : 00–9:00.

The digits in Figure 8 are labeled according to the order of arrival of the patients. Due to the characteristic of large scale, we only give some scheduling results in each time period. Interestingly, the patient service order based on reinforcement learning is almost completely different from that in the FCFS rules, and this is because of the fact that the reinforcement learning algorithm can achieve a better trade-off between the waiting time of patients and the setup cost of doctors, and it also takes into account the service rate of doctors, the different medical costs from doctors, and the personal choice preference of some patients.

Our experimental results can provide decision support for managers. In a queuing system with multiple types of patients and multiple service desks, we can reasonably arrange the number of patients and service order of each service desk according to the choice preference of patients. This has certain reference significance for other organizations with scarce resources. According to experiments conducted on the computer using Python, we can obtain satisfactory results within 3 minutes through the Q-learning. Due to the large scale of the proposed problem, the relatively optimal results that can be obtained within a few minutes are enough to prove that our proposed method has high effectiveness for solving the dynamic scheduling problem with multiple types of patients and multiple service queues.

6. Conclusions and Future Research

In this paper, a dynamic server scheduling problem in a special in vitro fertilization (IVF) queuing network, which is developed in an integrated cloud healthcare system, is investigated in order to address the prolonged waiting problem for IVF service. Based on continuous-time Markov procedure, a mathematical model is established, in which multiple types of patient selection preferences and multiple doctors with different service rates are considered simultaneously. To solve this model, a Q-learning-based solution method is proposed, where the reward functions are designed according to four conflicting cost functions: setup cost, waiting cost, penalty cost, and medical cost. A series of simulation experiments that are generated according to the actual data from Shenyang cloud hospital are carried out to validate the performance of the proposed reinforcement learning (RL) method for the investigated dynamic server scheduling problem in IVF queuing network.

The main contributions in this work lie in three aspects. Firstly, the IVF queueing network developed in cloud healthcare system is helpful to cope with the prolonged waiting problem of IVF medical service. Experimental results show that the waiting costs of patients decrease significantly in this integrated IVF queueing network. Secondly, dynamic server scheduling decision, that is, allocation of doctors in general hospital, is important to improve the efficiency of IVF medical service system. Our developed Markov model can exhibit a very nice flexibility in terms of both patient selection preference and medical resource utilization. Finally, the RL method is effective in solving the investigated problem in this paper. Based on the experimental results, our proposed Q-learning-based solution method significantly outperforms the traditional service rule. In general, the proposed methodology in this paper can not only make good use of the bottleneck medical resources in general hospital but also improve the utilization rate of idle medical resources in community hospital and it enables providing management insights for the cooperation of hospitals in hierarchical medical system.

This work can be extended in future researches. Firstly, we can consider the occupation of resources in the

telemedicine service process of IVF in community hospitals, as well as the influence of the success rate of surgery on the queuing system. Secondly, how to balance the interests of patients and multiple hospitals is also an interesting research issue. Finally, we can further analyze the impact of telemedicine service process in community hospitals on the service rate of doctors in general hospitals.

Data Availability

All data are included within the article.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] E. Mantikou, M. A. F. M. Youssef, M. van Wely et al., "Embryo culture media and IVF/ICSI success rates: a systematic review," *Human Reproduction Update*, vol. 19, no. 3, pp. 210–220, 2013.

[2] Y. Fu, H. Wang, J. Wang, and X. Pu, "Multiobjective modeling and optimization for scheduling a stochastic hybrid flow shop with maximizing processing quality and minimizing total tardiness," *IEEE Systems Journal*, vol. 99, pp. 1–12, 2020.

[3] W. J. Zhang and C. A. van Luttervelt, "Toward a resilient manufacturing system," *CIRP Annals*, vol. 60, no. 1, pp. 469–472, 2011.

[4] D. Ouelhadj and S. Petrovic, "A survey of dynamic scheduling in manufacturing systems," *Journal of Scheduling*, vol. 12, no. 4, p. 417, 2009.

[5] J. W. Wang, R. L. Dou, R. R. Muddada et al., "Management of a holistic supply chain network for proactive resilience: theory and case study," *Computers & Industrial Engineering*, vol. 125, pp. 668–677, 2017.

[6] Y. Zhang, S. Cheng, Y. Shi, D.-w. Gong, and X. Zhao, "Cost-sensitive feature selection using two-archive multi-objective artificial bee colony algorithm," *Expert Systems with Applications*, vol. 137, pp. 46–58, 2019.

[7] Y. Fu, D. Wu, Y. Wang, and H. Wang, "Facility location and capacity planning considering policy preference and uncertain demand under the One Belt One Road initiative," *Transportation Research Part A: Policy and Practice*, vol. 138, pp. 172–186, 2020.

[8] W. J. Zhang, J. W. Wang, and Y. Lin, "Integrated design and operation management for enterprise systems," *Enterprise Information Systems*, vol. 13, no. 4, pp. 424–429, 2019.

[9] H. Ouyang, N. T. Argon, and S. Ziya, "Allocation of intensive care unit beds in periods of high demand," *Operations Research*, vol. 68, no. 2, pp. 591–608, 2020.

[10] B. Denton and D. Gupta, "A sequential bounding approach for optimal appointment scheduling," *IIE Transactions*, vol. 35, no. 11, pp. 1003–1016, 2003.

[11] J. Huang, B. Carmeli, and A. Mandelbaum, "Control of patient flow in emergency departments, or multiclass queues with deadlines and feedback," *Operations Research*, vol. 63, no. 4, pp. 892–908, 2015.

[12] S. A. Erdogan, A. Gose, and B. T. Denton, "Online appointment sequencing and scheduling," *IIE Transactions*, vol. 47, no. 11, pp. 1267–1286, 2015.

[13] S. He, M. Sim, and M. Zhang, "Data-driven patient scheduling in emergency departments: a hybrid robust-stochastic approach," *Management Science*, vol. 65, no. 9, pp. 4123–4140, 2019.

[14] X. Li, J. Wang, and R. Y. K. Fung, "Approximate dynamic programming approaches for appointment scheduling with patient preferences," *Artificial Intelligence in Medicine*, vol. 85, pp. 16–25, 2018.

[15] D. Wang, K. Muthuraman, and D. Morrice, "Coordinated patient Appointment scheduling for a multistation healthcare network," *Operations Research*, vol. 67, no. 3, pp. 599–618, 2019.

[16] R. Atar, A. Goswami, and A. Shwartz, "Risk-sensitive control for the parallel server model," *SIAM Journal on Control and Optimization*, vol. 51, no. 6, pp. 4363–4386, 2013.

[17] N. Liu, S. R. Finkelstein, M. E. Kruk, and D. Rosenthal, "When waiting to see a doctor is less irritating: understanding patient preferences and choice behavior in appointment scheduling," *Management Science*, vol. 64, no. 5, pp. 1975–1996, 2017.

[18] A. K. Dogru and S. H. Melouk, "Adaptive appointment scheduling for patient-centered medical homes," *Omega*, vol. 85, pp. 166–181, 2019.

[19] H.-J. Schütz and R. Kolisch, "Approximate dynamic programming for capacity allocation in the service industry," *European Journal of Operational Research*, vol. 218, no. 1, pp. 239–250, 2012.

[20] V.-A. Truong, "Optimal advance scheduling," *Management Science*, vol. 61, no. 7, pp. 1584–1597, 2015.

[21] R. Hassin and L. Ravner, "Delay-minimizing capacity allocation in an infinite server-queueing system," *Stochastic Systems*, vol. 9, no. 1, pp. 27–46, 2019.

[22] T. R. Rohleder, D. Cooke, P. Rogers, and J. Egginton, "Coordinating health services: an operations management perspective," *Handbook of Healthcare Operations Management*, vol. 184, pp. 421–445, 2013.

[23] S. Saghafian, W. J. Hopp, S. Iravani et al., "Workload management in telemedical physician triage and other knowledge-based service systems," *Management Science*, vol. 64, no. 11, pp. 4967–5460, 2018.

[24] B. Rajan, T. Tezcan, and A. Seidmann, *The Process Implications of Using Telemedicine for Chronically Ill Patients: Analyzing Key Consequences for Patients and Medical Sspecialists*, California State University, Hayward, CA, USA, 2015.

[25] T. H. Westerdale, "An approach to credit assignment in classifier systems," *Complexity*, vol. 4, no. 2, pp. 49–52, 2015.

[26] S. Chen, Q. Qin, J. F. Chen, and Y. H. Shi, "Brain storm optimization algorithm: a review," *Artificial Intelligence Review*, vol. 46, no. 4, pp. 445–458, 2016.

[27] Q. Qin, S. Cheng, Q. Zhang, L. Li, and Y. Shi, "Particle swarm optimization with interswarm interactive learning strategy,"

IEEE Transactions on Cybernetics, vol. 46, no. 10, pp. 2238–2251, 2016.

[28] H. Lu, R. R. Zhou, S. Cheng, and Y. H. Shi, "Multi-center variable-scale search algorithm for combinatorial optimization problems with the multimodal property," *Applied Soft Computing*, vol. 84, 2019.

[29] Q. Qin, S. Cheng, X. Chu, X. Lei, and Y. Shi, "Solving non-convex/non-smooth economic load dispatch problems via an enhanced particle swarm optimization," *Applied Soft Computing*, vol. 59, pp. 229–242, 2017.

[30] Z. Huang, W. M. P. van der Aalst, X. Lu, and H. Duan, "Reinforcement learning based resource allocation in business process management," *Data & Knowledge Engineering*, vol. 70, no. 1, pp. 127–145, 2011.

[31] D. B. Noureddine, A. Gharbi, and S. B. Ahmed, "Multi-agent deep reinforcement learning for task allocation in dynamic environment," in *Proceedings of the 12th International Conference on Software Technologies*, pp. 17–26, Madrid, Spain, 2017.

[32] Z. Xiao, S. X. Ma, and S. Y. Zhang, "Learning task allocation for multiple flows in multi-agent systems," in *Proceedings of the 2009 International Conference on Communication Software and Networks*, pp. 153–157, Chengdu Sichuan, China, 2009.

[33] A. Asghari, M. K. Sohrabi, and F. Yaghmaee, "Online scheduling of dependent tasks of cloud's workflows to enhance resource utilization and reduce the makespan using multiple reinforcement learning-based agents," *Soft Computing*, vol. 24, no. 21, pp. 16177–16199, 2020.

[34] T. Wauters, K. Verbeeck, P. De Causmaecker, and G. Vanden Berghe, "A learning-based optimization approach to multi-project scheduling," *Journal of Scheduling*, vol. 18, no. 1, pp. 61–74, 2015.

[35] R. S. Sutton and G. A. Barto, *Reinforcement learning: an introduction*, MIT Press, vol. 17, no. 2, p. 322, Cambridge, MA, 1998.

[36] Y. Gao, S. F. Chen, and X. Lu, "Research on reinforcement learning technology: a review," *Acta Automatica Sinica*, vol. 30, no. 1, pp. 86–100, 2004.