International Journal of Applied Optimization Studies
Year 2019 Volume 02, Issue 03, pp. 34-48.
www.ijaos.com

Optimal Number of Service Providers in a Queuing System

Saeid Delshad [a,*], **Sajjad Ebadi Torkayesh** [b]

[a] *Department of Industrial Engineering, Clemson University, Clemson, South Carolina, USA*
[b] *Department of Industrial Engineering, Islamic Azad University of Tabriz, Tabriz, Iran*

Abstract

Manufacturing and service organizations make use of the queuing theory, not only to reduce the waiting time of their customers for optimal decision-making while determining the level of resources required for investment but also to meet the customers' satisfaction as much as possible. This is particularly important in the economic sustainability of companies in competitive environments. In most cases, companies face some troubles in determining the number of staffs. In this regard, the current study proposes different solutions to different potential problems. These solutions include the use of the coefficient of productivity, Markov systems, non-Markov systems and simulation tools. All the solutions presented in this paper have been used in a case study for a company in the service sector. The results show that our system can respond to the customers using five service providers such that customers would stay in the system less than or equal to one day. The results are also relatively close to each other which indicate that we can use each of these solutions to determine the number of service providers.

Original Article:

Received 2018-11-27
Revised 2019-04-20
Accepted 2019-05-05

Keywords:
Queuing theory;
Service providers;
Coefficient of productivity;
Markov systems;
Simulation.

1. Introduction

Although waiting in line is unpleasant; unfortunately, it is an inevitable reality of life. People face a variety of queues in their daily life that lead to loss of time, energy, and money. Eliminating the unfavorable consequences of waiting in a queue is not possible without knowing the characteristics of this phenomenon. Queuing theory deals with the study of queues from a mathematical point of view, studies the effects of constituting factors of the queue, and the logical methods to reduce waiting time. Although a queue cannot be eliminated, the wastes resulting from it can be reduced as much as possible.

Most of the organizations act based on either traditional or decision-making methods to determine the number of needed staffs. The first one, due to the non-scientific nature and the second one because of its approximate nature are not suitable for this purpose. Currently, the queuing theory as a valuable tool is applied in various professions. In this research, we introduce four methods of queuing theory for determining the number of personnel and we

*Corresponding author email address: sdelsha@clemson.edu

indicate that queuing theory can be used as a scientific and reliable method to determine the number of staff. Finally, we make a case example for all presented solutions.

One of the main activities of service managers is to determine the number of personnel and resources in such a way that neither any part of the organization end up unemployed nor customers spend too much time in queues before receiving services, considering customers may leave organizations because of long queues. The purpose of this research is to offer different solutions for determining the number of personnel in organizations using queuing theory.

Among various optimization problems and with exact or approximation methods (Tirkolaee et al., 2019; Goli et al., 2019), the queuing theory methodology is known an appropriate approach for modeling complex systems. In the literature, there are different application of this approach for problems, such as: healthcare systems (Malmir et al., 2016; Jahantigh et al., 2016; Aghighi and Malmir, 2016), customer relationship management (Tabandeh and Bastan, 2014), and Quality Control systems (Hashemian et. al., 2016).

Queuing theory has originated from a Danish engineer, Erlang, who studied the factors and relationships in the telephone system by testing increasing and decreasing demand (Abramowitz and Stegun, 1965). Eight years later, he published detailed studies on automation of telephone system, and the results of existent relationships that was the basis of his queuing theory. At the end of World War II, he developed quickly the usage of queuing models in the public and business domains (Gross, 2008).

Antoniol et al. (2004) conducted a study in which they presented queuing theory and stochastic simulation as a useful means to assess staff, evaluate the service level and review the probable deadline for the project. It was discovered later that it provides an efficient tool for managers provided that it is complemented by other managerial tools. It can also be used in order to prioritize activities, prevent conflicts and check the availability of the resources used. Lan et al. (2005) addressed the optimal allocation of human resources with limited server and finite capacity of the queue at a food store. They not only have presented the mathematical model for the queuing theory with limited server and finite capacity of queue, but also have reviewed an objective function considering the mental attitude of customers leaving, and finally, they proposed a step by step algorithm to optimize human resource allocation.

Fomundam and Herrmann (2007) introduced a wide range of queuing theory results in multiple fields, e.g., waiting time, utilization analysis, system design and appointment system, which are implemented at various scales, including individual group (unit), health care sector, and the regional health care systems.

In contrast to the conventional thinking which Physicians' time is more valuable than the patient's time, Mardiah and Basri (2013) examined the time of both groups as equally significant. The results of this research developed a good strategy for improving the appointment system in such a way that the resources and capacity of the hospital will efficiently.

Ameh et al. (2013) examined the level of patients' satisfaction in a Nigerian hospital using queuing theory. The results of their research showed that patients who are less than an hour waiting in line, are more satisfied so they suggest that the hospital would increase customer satisfaction through increasing the number of physicians.

Since the phenomenon of waiting in line is an unenjoyable experience for customers, and this phenomenon can also be costly for service organizations, one of the main activities of the managers is reducing the waiting time of customers in the service queue. In this research, we want to determine the number of service personnel required to engage in a service organization using the concepts from queuing theory, so that the customers' waiting time in the queue would not exceed the standard value. To determine the number of service personnel, organizations are often faced with the problem of lack of information, and also determining the type of queuing system is often difficult for them. In this study, we examine four different solutions to determine the number of servers. The first method, namely the coefficient of productivity can be used when the type of the system is not clear and the only information available to us is the arrival rate and service rate. The second approach which is the Markov systems can be used when the system is fully recognized to us and we have all the required information available. The third method which is the non-Markov systems can be used when we are relatively aware of the type of the system. For example, we know that the arrival rate follows Poisson process, but we are not sure what type of process the service rate follows. The fourth method namely simulation can be used in most cases and usually when the system is large. This last method can provide more reliable answers in complicated systems.

2. Model

Queuing theory is one of the oldest and most developed techniques used in the analysis of waiting lines. Queue, in its simplest form, is where the recipient of the service waits until his turn and the queuing theory is associated primarily with systems in which customer's arrival is random as well as serving them. Queuing systems can be studied with respect to their two main features; customer arrival rate to system and customer exit from the system (Russell and Taylor, 2003). In general, a queuing system can be shown schematically as Figure 1.

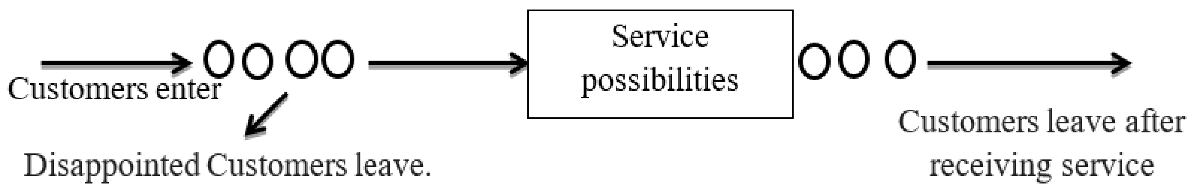

Fig. 1. Schematic diagram of a Queue process.

According to Jacobs et al. (2004) a queuing system essentially consists of three components:

1. The population of clients entered to the system that can be considered infinite or finite.
2. The service provider system that includes the queue and several servers. The queue length, the number of queues and queue laws are some of the features of waiting queue.
3. The exit status of the customers meaning whether the recipients of service will return to the community or not.

In the literature on queuing theory and queuing systems, the queuing system is considered as a static system in which the rate of customer arrival and providing service in the system is constant over time. These assumptions can be largely unrealistic, as depending on the length of the queue and the work pressure, the rate of service can change over time.

To measure the performance of a queuing system, evaluation criteria are mainly based on the following three variables formulated through equations:

1. The queue length: the number of customers in the queue waiting for service or the number of customers in the system.
2. The waiting time for each customer in queue or system. It should be noted that the waiting time in the system is the sum of waiting time in the queue and the time of customers wait for receiving the service.
3. The percentage of time that the system is idle due to lack of customer (or the percentage of time that the system is working).

We should bear in mind that in most systems, these variables have a random nature. Consequently, the evaluation criteria of the system are generally the expected value of these random variables. It should be noted that other criteria could also be used if needed. In this study, we examine the waiting time or average time of customer's presence in the system so that accordingly, we could determine the number of service providers.

The required evaluation criteria of a system in long-term using the Little Law are obtained as follows. These equations that are particularly important in queuing theory are true about all types of queuing systems:

$$L = \lambda W \tag{1}$$

$$L_q = \lambda W_q \tag{2}$$

$$W = W_q + \frac{1}{\mu} \tag{3}$$

3. Determining the Number of Service Providers

To measure the performance of a queuing system, evaluation criteria are mainly based on the following three variables: In this study, we consider the number of customers in the system as the system state, and therefore, we want to specify the number of servers in such a way that the waiting time in the system reaches a certain level. We can use four different solutions to do this as follows.

3.1. Coefficient of Productivity

A system is in a stable state when $\rho < 1$ and the ρ value can be obtained from equation (4). Now, we want to use the waiting time for deciding about the system. If we want to reduce the waiting time and reach the specified level, we should increase the number of servers, i.e. $W \leq x$ .Then we calculated the L_q value of a system that represents whole community which in this research can be obtained by Eq. (5), then the W value that we are looking for can be obtained from Eqs. (2)-(3).

$$\rho = \frac{\lambda}{\mu m} \tag{4}$$

$$L_q = \frac{1+r}{2r} \times \frac{\rho^2}{1-\rho} \tag{5}$$

Since λ and μ values (μ per server) are constant, the W value can be decreased as much as x unit or less by determining the L_q for different values of *m*.

3.2. Markov Systems

In this method, we must detect the type of queuing system that in this research is $M/E_r/m$, then using the equation of the average time spent in the system in long-term (*W*), it can be obtained for different amounts of *m* and whenever *W* was less than the requested amount, the number of servers is optimal.

For example, in the $M/E_r/1$ system, the W_q value can be obtained from equation (6).

$$W_q = \frac{r+1}{2r} \times \frac{\lambda}{\mu(\mu-\lambda)} \tag{6}$$

3.3. Non-Markov Systems

In this method, we calculate the *W* value for different *m* values in the $G/G/m$ queuing system, and whenever *W* value was less than the determined amount, it has reached the optimal solution.

For example, the equations for the $M/G/m$ queuing system can be written as follows:

$$L_q = \frac{r^{m-1}\left[\lambda^2 Var(s) + r^2\right]}{2(m-1)!(m-r)^2\left[\sum_{j=0}^{m-1}\frac{r^j}{j!} + \frac{r^m}{(m-1)!(m-r)}\right]} \tag{7}$$

$$r = \frac{\lambda}{\mu} \tag{8}$$

where W_q value can be calculated from Eq. (2) and then the *W* value can be calculated using the equation (3).

3.4. Simulation

Usually, if a system could be modeled mathematically, it will not be simulated, but when the system is complex and includes a lot of interactive elements, mathematical analysis of the system may be beyond the abilities of the analyst. However, since simulation is able to analyze such a system, complexity is not an unsolvable issue. So, the most important benefit of simulation is that it can be used for a wide variety of problems, including complex systems which their mathematical modeling is not suitable, and also where successful construction of a simulation model requires less experience in mathematics and probability theory than the knowledge required for a similar analysis by mathematical modeling. Thus, the ease of using simulation can be considered as one of the positive points that can also be detrimental in some situations, i.e., some researchers prefer using simulation even where mathematical techniques are easy to use. The significant problem of using the simulation for system analysis is the inherent computational costs associated with the long-term use of a computer.

For a queuing system simulation, we draw the general flowchart (Figure 2) and then determine the target value using the existent information. For example, in this study, we have the data of customer arrival and service rates and we want to determine the number of servers in such a way that the customers' presence in the system is less than or equal to one day.

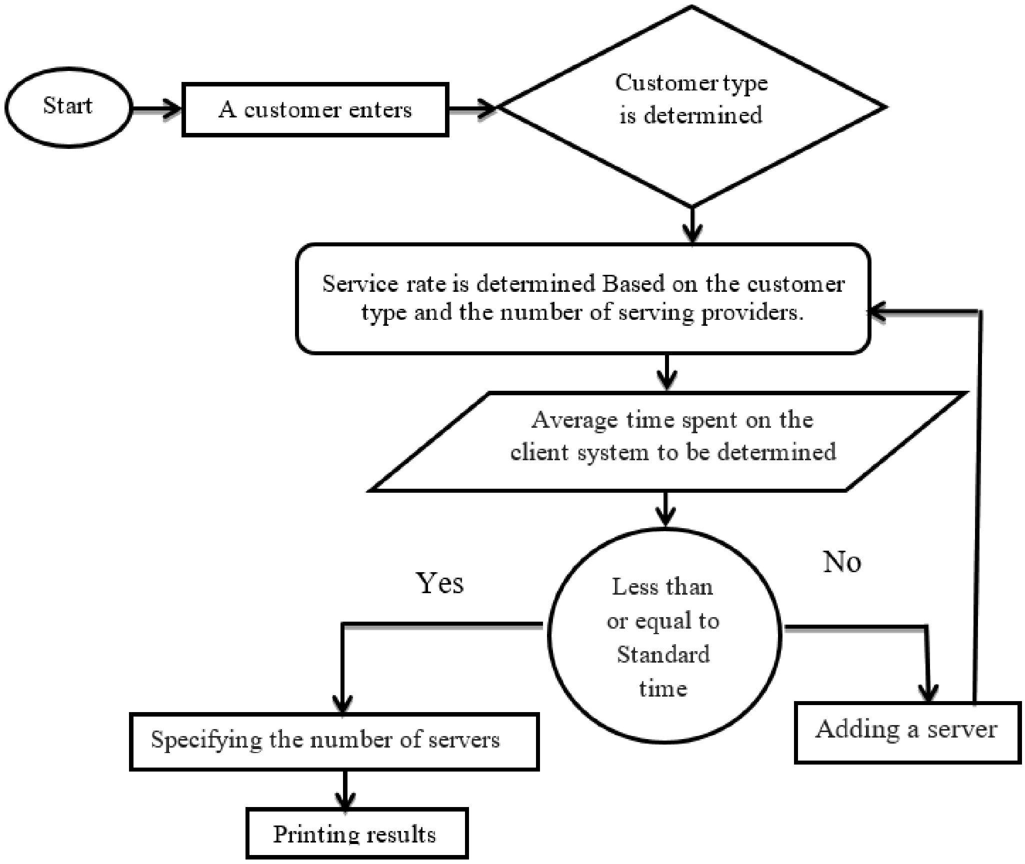

Fig. 2. Simulation flowchart.

4. Case Study

In this paper, we study an engineering services company that is involved in the testing of petroleum products. We are going to determine the number of testing personnel in such a way that the answer of each sample be determined and sent within a period of one day after delivering samples to a lab (Given that the samples are petroleum products and hence highly flammable, Therefore, the answering time, whether approval or disapproval, should not take more than two days, assuming that delivering the sample to the laboratory usually take a day). Testing each sample, from beginning to end should only be performed by an examiner to have traceability in case of an event. The incoming samples to the company are composed of different types including foots oil and furfural, pitch, other oils, paraffin, solvents, lubricants, etc. generally the incoming samples can be categorized, based on the different stages of the test, into 6 different groups. The useful working hours per day is 5 hours and 26 days per month. Incoming samples to the lab follow the Poisson process with an arrival rate of 57 samples per month, the number of samples in each group is presented in Table 1. Response time and the number of test stages of each group of samples are presented in Table 2. According to Table 2, in this case study, we deal with six different queuing systems that are given in Table 3. Now, to determine the number of servers, we use four different methods mentioned above:

Table 1. Number of samples of each group.

Number	Incoming samples	Size
1	foots oil and furfural	8
2	Pitch	4
3	Paraffin, motor oil, base oil, gear oil, etc.	31
4	Gasoline, Paint thinner, solvents, hydrocarbons, etc.	9
5	Grease	2
6	Ethylene glycol	3

Table 2. Number of the test stages and response time of samples.

Number	Incoming samples	Number of test stages	Response time of a sample (in minutes)
1	foots oil and furfural	3	140
2	Pitch	1	10
3	Paraffin, motor oil, base oil, gear oil, etc.	9	410
4	Gasoline, Paint thinner, solvents, hydrocarbons, etc.	6	250
5	Grease	6	290
6	Ethylene glycol	6	190

Table 3. Queuing systems existing in the lab.

Number	Incoming samples	Types of queuing systems
1	foots oil and furfural	$M/E_3/m$
2	Pitch	$M/E_1/m$ or $M/M/m$
3	Paraffin, motor oil, base oil, gear oil, etc.	$M/E_9/m$
4	Gasoline, Paint thinner, solvents, hydrocarbons, etc.	$M/E_6/m$
5	Grease	$M/E_6/m$
6	Ethylene glycol	$M/E_6/m$

4.1. Using the Coefficient of Productivity

$$\lambda = 57\frac{sample}{month} = 2.19\frac{sample}{day} \tag{9}$$

The μ value will vary for each type of sample, so that the average value per server is equal to:

$$\mu = 1.1\frac{sample}{day} \tag{10}$$

As a result, the coefficient of productivity is equal to:

$$\rho = \frac{\lambda}{m\mu} \Rightarrow m = 1 \Rightarrow \rho = \frac{2.19}{1.1} \Rightarrow \rho = 1.99 > 1 \tag{11}$$

Because with a service provider, the system isn't in the stable state, we don't continue the calculations for a service provider and checking system with two servers.

$$m = 2 \Rightarrow \rho = \frac{2.19}{2(1.1)} \Rightarrow \rho = 0.99 < 1 \tag{12}$$

$$L_q = \frac{1+r}{2r} \times \frac{\rho^2}{1-\rho} \Rightarrow L_q = \frac{1+6}{2(6)} \times \frac{0.99^2}{1-0.99} \Rightarrow L_q = 0.58 \times 98 \Rightarrow L_q = 56.84 \tag{13}$$

$$W_q = \frac{1}{\lambda} L_q \Rightarrow W_q = \frac{1}{2.19} \times 56.84 \Rightarrow W_q = 25.59 \tag{14}$$

$$W = W_q + \frac{1}{\mu} \Rightarrow W = 25.59 + \frac{1}{1.1} \Rightarrow W = 26.49 \tag{15}$$

The waiting time is longer than the specified time, so check the waiting time with 3 servers:

$$m = 3 \Rightarrow \rho = \frac{2.19}{3(1.1)} \Rightarrow \rho = 0.66 < 1 \tag{16}$$

$$L_q = \frac{1+r}{2r} \times \frac{\rho^2}{1-\rho} \Rightarrow L_q = \frac{1+6}{2(6)} \times \frac{0.66^2}{1-0.66} \Rightarrow L_q = 0.58 \times 1.26 \Rightarrow L_q = 0.73 \tag{17}$$

$$W_q = \frac{1}{\lambda} L_q \Rightarrow W_q = \frac{1}{2.19} \times 0.73 \Rightarrow W_q = 0.33 \tag{18}$$

$$W = W_q + \frac{1}{\mu} \Rightarrow W = 0.33 + \frac{1}{1.1} \Rightarrow W = 1.23 \tag{19}$$

The waiting time is greater than the standard value, so we try again with 4 servers:

$$m = 4 \Rightarrow \rho = \frac{2.19}{4(1.1)} \Rightarrow \rho = 0.498 < 1 \tag{20}$$

$$L_q = \frac{1+r}{2r} \times \frac{\rho^2}{1-\rho} \Rightarrow L_q = \frac{1+6}{2(6)} \times \frac{0.498^2}{1-0.498} \Rightarrow L_q = 0.58 \times 0.49 \Rightarrow L_q = 0.2842 \tag{21}$$

$$W_q = \frac{1}{\lambda} L_q \Rightarrow W_q = \frac{1}{2.19} \times 0.2842 \Rightarrow W_q = 0.1297 \tag{22}$$

$$W = W_q + \frac{1}{\mu} \Rightarrow W = 0.1297 + \frac{1}{1.1} \Rightarrow W = 1.03 \tag{23}$$

The average spent time by a customer in the company is not less than waiting time, so examine the problem with 5 servers:

$$m = 5 \Rightarrow \rho = \frac{2.19}{5(1.1)} \Rightarrow \rho = 0.398 < 1 \tag{24}$$

$$L_q = \frac{1+r}{2r} \times \frac{\rho^2}{1-\rho} \Rightarrow L_q = \frac{1+6}{2(6)} \times \frac{0.398^2}{1-0.398} \Rightarrow L_q = 0.58 \times 0.26 \Rightarrow L_q = 0.1508 \tag{25}$$

$$W_q = \frac{1}{\lambda} L_q \Rightarrow W_q = \frac{1}{2.19} \times 0.1508 \Rightarrow W_q = 0.0688 \tag{26}$$

$$W = W_q + \frac{1}{\mu} \Rightarrow W = 0.0688 + \frac{1}{1.1} \Rightarrow W = 0.9778 \tag{27}$$

Since with increasing the number of service providers, the waiting time in line is reduced, if we use 5 servers, L_q value will also decrease and therefore the W value reaches the standard level, so the number of servers with this method is equal to 5.

4.2. Using Markov Systems

Here, we calculate the W value using Eqs. (4)-(5). As there are different $M / E_r / m$ systems in this study, we calculated the value of each system for different numbers of service providers separately. The results are indicated in Tables 4-8.

$$W_q = \frac{r+1}{2r} \times \frac{\lambda}{\mu(\mu - \lambda)} \tag{28}$$

$$W = W_q + \frac{1}{\mu} \tag{29}$$

Due to extensive computations, only the results are indicated in Tables 4 to 8. Given that the total value is greater than the specified standard value, we examine the average waiting time for two service providers in Table 5. Since the obtained total value is again greater than the standard value, we examine the waiting time with 3 service providers in Table 6.

Considering that the average waiting time with three servers is higher than the standard time (one day) required, therefore 3 servers is not enough, so the case with 4 servers is examined in Table 7. Again, the waiting time exceeds the determined value, so we address reviewing five servers in Table 8. The customer's average waiting time in the system with 5 servers is less than the determined value, so 5 servers are sufficient.

Table 4. Determining the W and W_q values per service provider.

Number	Incoming samples	r	λ	μ	W_q	W
1	foots oil and furfural	3	0.31	2.14	0.052	0.52
2	Pitch	1	0.15	30.0	0.00017	0.03
3	Paraffin, motor oil, base oil, gear oil, etc.	9	1.19	0.73	1.985	3.35
4	Gasoline, Paint thinner, solvents, hydrocarbons, etc.	6	0.35	1.20	0.197	1.03
5	Grease	6	0.08	1.03	0.047	1.018
6	Ethylene glycol	6	0.11	1.58	0.027	0.66
Total						6.608

Table 5. Determining W and W_q values for two service providers.

Number	Incoming samples	r	λ	μ	W_q	W
1	foots oil and furfural	3	0.31	4.28	0.01188	0.24548
2	Pitch	1	0.15	60.00	0.000417	0.01708
3	Paraffin, motor oil, base oil, gear oil, etc.	9	1.19	1.46	1.6775	2.36243
4	Gasoline, Paint thinner, solvents, hydrocarbons, etc.	6	0.35	2.40	0.041238	0.82904
5	Grease	6	0.08	2.06	0.01137	0.49680
6	Ethylene glycol	6	0.11	3.16	0.00662	0.32307
Total						4.2739

Table 6. Determining W and W_q values for three service providers.

Number	Incoming samples	r	λ	μ	W_q	W
1	foots oil and furfural	3	0.31	6.42	0.00521	0.16097
2	Pitch	1	0.15	90.00	0.0000185	0.01113
3	Paraffin, motor oil, base oil, gear oil, etc.	9	1.19	2.19	0.29865	0.75527
4	Gasoline, Paint thinner, solvents, hydrocarbons, etc.	6	0.35	3.60	0.01734	0.29511
5	Grease	6	0.08	3.09	0.00498	0.32860
6	Ethylene glycol	6	0.11	4.74	0.00290	0.21387
Total						1.76495

4.3. Using Non-Markov Systems

We calculate the W value using Eqs. (6)-(8) for different amounts of m and whenever the W value would be less than the standard value, the required number of servers is obtained.

Here, since the arrival of the samples follows a Poisson process, the variance is equal to:

$$Var(s) = \frac{1}{\lambda^2} \Rightarrow Var(s) = \frac{1}{(2.19)^2} \Rightarrow Var(s) = 0.2085 \tag{30}$$

Table 7. Determining W and W_q values for four service providers.

Number	Incoming samples	r	λ	μ	W_q	W
1	foots oil and furfural	3	0.31	8.56	0.00289	0.11971
2	Pitch	1	0.15	120.0	0.0000104	0.00834
3	Paraffin, motor oil, base oil, gear oil, etc.	9	1.19	2.92	0.12956	0.47202
4	Gasoline, Paint thinner, solvents, hydrocarbons, etc.	6	0.35	4.80	0.00950	0.21783
5	Grease	6	0.08	4.12	0.00279	0.24551
6	Ethylene glycol	6	0.11	6.32	0.00162	0.15823
Total						1.22164

Table 8. Determining W and W_q values for five service providers.

Number	Incoming samples	r	λ	μ	W_q	W
1	foots oil and furfural	3	0.31	10.7	0.00278	0.09623
2	Pitch	1	0.15	150.0	0.0000067	0.00667
3	Paraffin, motor oil, base oil, gear oil, etc.	9	1.19	3.65	0.07287	0.34684
4	Gasoline, Paint thinner, solvents, hydrocarbons, etc.	6	0.35	6.00	0.00598	0.17264
5	Grease	6	0.08	5.15	0.00177	0.19594
6	Ethylene glycol	6	0.11	7.90	0.00103	0.12761
Total						0.94593

For a single service provider $(m = 1)$ the values for L_q, W_q and W are equal to:

$$r = \frac{\lambda}{\mu} \Rightarrow r = \frac{2.19}{1.1} \Rightarrow r = 1.99 \tag{31}$$

$$L_q = \frac{r^{m-1}\left[\lambda^2 Var(s) + r^2\right]}{2(m-1)!(m-r)^2\left[\sum_{j=0}^{m-1}\frac{r^j}{j!} + \frac{r^m}{(m-1)!(m \times r)}\right]} \tag{32}$$

$$L_q = \frac{1.99^{1-1}\left[2.19^2 \times 0.2085 + 1.99^2\right]}{2(1-1)!(1-1.99)^2\left[\sum_{j=0}^{1-1}\frac{1.99^j}{j!} + \frac{1.99^1}{(1-1)!(1 \times 1.99)}\right]} \Rightarrow L_q = \frac{4.4167}{3.9204} \Rightarrow L_q = 1.13 \tag{33}$$

$$W_q = \left(\frac{1}{\lambda}\right)L_q \Rightarrow W_q = \left(\frac{1}{2.19}\right)1.13 \Rightarrow Wq = 0.5198 \tag{34}$$

$$W = W_q + \frac{1}{\mu} \Rightarrow W = 0.5198 + \frac{1}{1.1} \Rightarrow W = 1.4288 \tag{35}$$

The obtained W value is higher than the standard time (one day), so we solve the problem with two servers again:

$$L_q = \frac{1.99^{2-1}\left[2.19^2 \times 0.2085 + 1.99^2\right]}{2(2-1)!(2-1.99)^2\left[\sum_{j=0}^{2-1}\frac{1.99^j}{j!} + \frac{1.99^2}{(2-1)!(2\times 1.99)}\right]} \Rightarrow L_q = \frac{9.87057}{0.00059} \Rightarrow L_q = 16533.61 \quad (36)$$

$$W_q = \left(\frac{1}{\lambda}\right)L_q \Rightarrow W_q = \left(\frac{1}{2.19}\right)16533.61 \Rightarrow W_q = 0.46 \times 16533.61 \Rightarrow W_q = 7605.46 \quad (37)$$

$$W = W_q + \frac{1}{\mu} \Rightarrow W = 7605.46 + \frac{1}{1.1} \Rightarrow W = 7606.36 \quad (38)$$

Again, the obtained W value is higher than the standard amount (one day), so we solve the problem with three service providers:

$$L_q = \frac{1.99^{3-1}\left[2.19^2 \times 0.2085 + 1.99^2\right]}{2(3-1)!(3-1.99)^2\left[\sum_{j=0}^{3-1}\frac{1.99^j}{j!} + \frac{1.99^3}{(3-1)!(3\times 1.99)}\right]} \Rightarrow L_q = \frac{19.6424}{18.89225} \Rightarrow L_q = 1.03970 \quad (37)$$

$$W_q = \left(\frac{1}{\lambda}\right)L_q \Rightarrow W_q = \left(\frac{1}{2.19}\right)1.03970 \Rightarrow W_q = 0.47474 \quad (38)$$

$$W = W_q + \frac{1}{\mu} \Rightarrow W = 0.47474 + \frac{1}{1.1} \Rightarrow W = 1.38383 \quad (39)$$

The average amount of time spent in the system (W) is higher than the specified time, so we conclude that 3 servers would not be sufficient, so we check it with four servers.

$$L_q = \frac{1.99^{4-1}\left[2.19^2 \times 0.2085 + 1.99^2\right]}{2(4-1)!(4-1.99)^2\left[\sum_{j=0}^{4-1}\frac{1.99^j}{j!} + \frac{1.99^4}{(4-1)!(4\times 1.99)}\right]} \Rightarrow L_q = \frac{39.08835}{272.55306} \Rightarrow L_q = 0.14341 \quad (40)$$

$$W_q = \left(\frac{1}{\lambda}\right)L_q \Rightarrow W_q = \left(\frac{1}{2.19}\right)0.14341 \Rightarrow W_q = 0.065483 \quad (41)$$

$$W = W_q + \frac{1}{\mu} \Rightarrow W = 0.065483 + \frac{1}{1.1} \Rightarrow W = 0.974573 \quad (42)$$

The result obtained is less than the standard value, so 4 servers is sufficient. In this method, because the system is considered as an approximate system, the value of 4 was obtained which is close to the solution obtained by the other methods.

4.4. Using Simulation

According to the information and the overall simulation flowchart for the study, we examine the given system with a simulation method which its results are presented in Table 9.

Table 9. The results of obtained by system simulation.

Number of servers	Arrival rate of customers in a day (λ)	Rate of serving the customer in a day (μ)	Average waiting time of customers in the system a day (W)
1	2.19	1.1	6.608
2	2.19	2.2	4.2739
3	2.19	3.3	1.76495
4	2.19	4.4	1.22164
5	2.19	5.5	0.94593

According to the overall simulation flowchart, the system is able to answer all of the incoming samples with 5 servers in less than a day.

5. Conclusion

In this study, we sought to determine the number of servers for an engineering services company active in the field of petroleum products to reduce its respond time of the incoming samples to less than a day. We have four different ways to do this. In the first method, we studied the system using the coefficient of productivity. According to the calculations, five service providers were known sufficient. The second method uses a Markov system to check the system. In this method, the number of 5 was also detected as the appropriate number of servers. But the main challenge of this method is the extensive amount of calculations that increases the chances of making mistakes in the calculation. In the third method, we examined the system using a non-Markov system by which 4 servers is considered appropriate that is one unit less than the actual value. The reason of this contradiction is the approximate nature of non-Markov systems. If we study the system using this method for 5 servers, we will observe that the results are very close to the other methods. In the last method, we examined the system using simulation of queuing systems. In this method, the 5 servers are also recognized as enough number of servers. It can be concluded that the system must operate with 5 servers so that any of the samples sent to the company would not stay in the company for more than one day. The results show that the presented methods are able to determine the number of personnel so that the customers' waiting time does not exceed the specified standard level.

In this study, we examined a system in terms of the average waiting time that customers spend in the queue. For a future study, one can examine queuing systems by changing one of the assumptions. Researchers can also compare two different systems using each or a combination of these four methods. Plus, one may examine a similar system with respect to other financial considerations and obtain the results accordingly.

References

Abramowitz, M., and Stegun, I. A. (1965). Handbook of mathematical functions: with formulas, graphs, and mathematical tables (Vol. 55). Courier Corporation.

Aghighi, A., and Malmir, B. (2016). Designing distribution networks of perishable products under stochastic demands and routs. In International conference on industrial engineering and operations management, Sep, Detroit, Michigan.

Almasi, F., and Malmir, B. (2015, March). Analyzing and Estimating the project success with a hybrid method based on fuzzy theory and decision-making analysis. In 2015 International Conference on Industrial Engineering and Operations Management (IEOM) (pp. 1-9). IEEE.

Ameh, N., Sabo, B., and Oyefabi, M. O. (2013). Application of queuing theory to patient satisfaction at a tertiary hospital in Nigeria. Nigerian medical journal: journal of the Nigeria Medical Association, 54(1), 64.

Antoniol, G., Cimitile, A., Di Lucca, G. A., and Di Penta, M. (2004). Assessing staffing needs for a software maintenance project through queuing simulation. IEEE Transactions on Software Engineering, 30(1), 43-58.

Babaee Tirkolaee, E., Abbasian, P., Soltani, M., and Ghaffarian, S. A. (2019). Developing an applied algorithm for multi-trip vehicle routing problem with time windows in urban waste collection: A case study. Waste Management and Research, 37(1_suppl), 4-13.

Bastan, M., Grösser N., S., Zadfallah, E. (2018). Model-Based Risk Assessment to Clinical Risk Management Policies. The 36th International Conference of the System Dynamics Society, Reykjavík, Iceland

Fomundam, S., and Herrmann, J. W. (2007). A survey of queuing theory applications in healthcare.

Goli, A., Tirkolaee, E. B., Malmir, B., Bian, G. B., and Sangaiah, A. K. (2019). A multi-objective invasive weed optimization algorithm for robust aggregate production planning under uncertain seasonal demand. Computing, 1-31.

Gross, D. (2008). Fundamentals of queuing theory. John Wiley and Sons.

Hashemian, S. M., Noorossana, R., Keyvandarian, A., & Shekary A, M. (2016). Performance of adaptive np-chart with estimated parameter. International Journal of Quality & Reliability Management, 33(6), 769-791.

Jacobs, F. R., Chase, R. B., and Aquilano, N. (2004). Operations management for competitive advantage. Boston: Mc-Graw Hill, 64, 70.

Jahantigh, F. F., and Malmir, B. (2016, March). A Hybrid Genetic Algorithm for Solving Facility Location Allocation Problem. In Proceedings of the 2016 International Conference on Industrial Engineering and Operations Management, Kuala Lumpur, Malaysia.

Lan, C. H., Lan, T. S., and Chen, M. S. (2005). Optimal human resource allocation with finite servers and queuing capacity. International journal of computer applications in technology, 24(3), 156-160.

Malmir, B., Dehghani, S., Jahantigh, F. F., and Najjartabar, M. (2016, June). A new model for supply chain quality management of hospital medical equipment through game theory. In

Proceedings of the 6th International Conference on Information Systems, Logistics and Supply Chain, ILS 2016.

Mardiah, F. P., and Basri, M. H. (2013). The Analysis of Appointment system to reduce outpatient waiting time at Indonesia's public hospital. Human Resource Management Research, 3(1), 27-33.

Russell, R. S., and Taylor, B. W. (2003). Operations management (Vol. 3). Upper Saddle River, NJ: Prentice Hall.

Tabandeh, M., and Bastan, M. (2014). Customers classification according to the grey-based decision-making approach and its application to bank queue systems. Asian Journal of Research in Banking and Finance, 4(7), 349-372.

Tirkolaee, E. B., Goli, A., Hematian, M., Sangaiah, A. K., and Han, T. (2019). Multi-objective multi-mode resource constrained project scheduling problem using Pareto-based algorithms. Computing, 1-24.