

## A steady-state analysis of a hair salon as a single-queue, multi-server system to optimize the waiting time in a queue

Sambhav Kharel<sup>a</sup>, Roshan Bhandari<sup>b</sup>, Satish K C<sup>c</sup>, Aayush Bhattra<sup>d\*</sup>

<sup>a,b,c,d</sup>*Department of Mechanical and Aerospace Engineering, Pulchowk Campus, IOE, Tribhuvan University, Pulchowk, Lalitpur, 44600, Nepal*

Received: 28 March 2020; Accepted: 09 May 2020; Published: 08 June 2020

---

### Abstract

The waiting lines and service systems are crucial parts of our daily life. Services like hospitals and clinics, banks, salons, restaurants, et cetera have a high influx of people during the working hours. Every system likes to avoid losing their desired customers due to a long wait in the queue. Increasing the number of servers obviously reduces waiting time, but is not the best solution since it adds expenses because the employees should be paid. On the other hand, not concerning about waiting time will result in customer dissatisfaction and ultimately the loss of valued customers. This article presents the analytical study of the overall time that a customer needs to spend on a hair salon for the current number of hairdressers as well as the impact of adding servers (hairdressers) on waiting time. It models the queuing system of the hair salon to optimize the waiting time for a customer who arrives at the salon. Further, it provides a basis to make a wise decision on adding servers. The system is modelled as a steady-state single queue, multi-server configuration system.

**Index Terms:** Probability Distribution, Waiting Line Model, Single Server Configuration, Kendall Notation

© 2020 Published by MECS Publisher. Selection and/or peer review under responsibility of the Research Association of Modern Education and Computer Science

---

\* Corresponding author.

E-mail address:

## 1. Introduction

Allocation Very often in our every-day life, we come across situations where we need to wait in a queue, wherever we go. Nobody likes waiting for a long time in a queue. To prevent this, reducing service time helps. But, depending on the nature of the job, it may not always be reduced. Some tasks such as registration for a medical check-up at the entrance, dressing hair at salons, etc. take nearly the same amount of time to serve each customer, which may not be always reduced below some value. So, the other alternative is increasing the number of servers. However, merely increasing the number of servers is not the finest solution since it results in increased expenses in the form of salary to be paid to the employee. On the other hand, if a customer sees a long queue, he may return without joining the queue due to which valued customers are lost by the service provider. Thus, service providers always seek to reduce the waiting time as far as possible with minimum expenses. This demands a careful analysis of the system to optimize the waiting time. This can be achieved by employing the queuing theory. Queuing theory is the mathematical analysis of waiting process in lines.

Queues form when the service provider has more demand but less service capacity. A basic queuing system consists of 3 processes: an arrival process, the queuing, and the service process provided to those customers. Most queuing problems focus on determining the level of service that a company should provide (Ragsdale, 2008).

A queuing system generally includes:

- Arrival process: The arrival process is the way of incoming customers. Customers may come alone or with somebody.
- Behavior: Some customers can wait for a long time whereas some customers may become impatient and leave. Sometimes, they may get irritated.
- Service to the customers: This includes the service time required for providing customer service, the number of service providers available, and the service discipline that is followed.
- Service discipline: It is the rule by which the service provider selects the next customer after serving one customer. Some have basic 'first-come first-service' rule whereas some prefer service based on sex, age group, and physical condition.

The report presented in this paper is a steady-state analysis of a hair salon which is modeled as the single-queue, multiple-server system which provides service to the customers on 'first-come-first-serve' basis.

The two major objectives of this research are to identify the optimal waiting time for a customer in a queue and hence to make decisions on the total number of servers to be deployed by the service provider (hair salon) as best possible solutions under the given circumstances.

The hair salon is located at Shankhamul, Kathmandu, the capital city of Nepal. It is a simple Nepalese-style salon owned by a local businessman.

The problem here is that, very often, customers have to wait for a long time for their turn and sometimes they even return by seeing the number of people waiting in the queue.

As an existing solution, two hairdressers (servers) are deployed. The limitation of the current solution is that the available two servers are not able to provide the service efficiently to the customers visiting the salon. On average, the customer needs to wait about 24 minutes in the queue before he gets his hair done.

To address this issue, a mathematical model, based on queuing theory, is developed and the analysis for the optimization of waiting time is carried out.

Salon centers are distributed almost evenly in this locality. This place, being the capital of the country, is densely populated and hence the customer flow is nearly the same every day. Therefore, a normal working day was selected for analysis.

The essential characteristics of the queuing process on this hair salon can be shown in a flow chart (Fig. 1.). Upon arrival, the customer can directly get service if any one of the servers is available. Otherwise, he has to

wait for his turn until one of the servers is free. He waits there for an average waiting time,  $W_q$ . Finally, the customer leaves the salon (system) after time  $W$ , on average, from arrival to completion of the service.

The analysis of the system is carried out as an M/M/s model. In the Kendall Notation above, the first ‘M’ refers to Markovian inter-arrival times (following an exponential distribution), second ‘M’ to Markovian service times (following an exponential distribution of service times) and ‘s’ refers to the number of servers available.

Hence, a notation M/M/1 denotes that the inter-arrival time of the customers follows the exponential distribution, the time taken by the server to serve each customer also follows the exponential distribution, and the total number of servers, i.e. the hair-dressers is two. However, those are the average arrival rate (since it is the steady-state analysis) and hence the average service duration that are employed in the analysis carried out in this case.

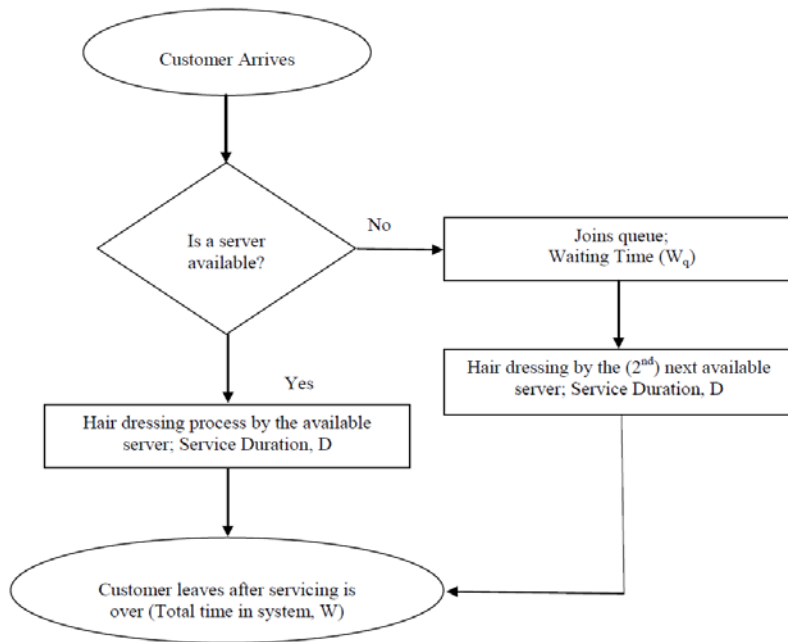


Fig.1. Flow diagram representing customer flow in a single-queue, multi-server configuration system (hair salon)

Nomenclature	
D	Service Duration
f	Frequency
$L_q$	Average no. of customers waiting for service
L	Average no. of customers in the system
N	Total number of customers visiting the hair salon per day
$P_0$	Probability that there are no customers in the system
$\lambda$	Average arrival rate
$\mu$	Average service rate
s	Number of servers
W	Overall-time spent by a customer in the system, on average
$W_q$	Average waiting time in a queue

## 2. Methodology

The data was collected for 8 hours from 9 am to 5 pm. The total number of customers arriving at the hairdresser ( $N$ ), the time they arrived, and the total time-duration ( $D$ ) for their service were noted from the footage of the CCTV camera placed at the hair salon. Then, the average arrival rate ( $\lambda$ ) of the customers is calculated using:

$$\lambda = \frac{N}{8} \text{ per hour} \quad (1)$$

From the frequency distribution, average service duration ( $D_{avg}$ ) and hence the average service rate ( $\mu$ ) are calculated as:

$$D_{avg} = \frac{\sum f \cdot D}{\sum f} \quad (2)$$

$$\mu = \frac{60}{D} \text{ per hour} \quad (3)$$

The arrival of the customers was random and hence the arrival rate is approximated using the Poisson random variable. The probability of arrival of  $x$  number of entities in a specified period (per hour in this case) is given by:

$$P(x) = \frac{\lambda \cdot e^{-x}}{x!} \quad (4)$$

where,  $x=0,1,2,3,\dots$

On the other hand, the inter-arrival time followed the exponential probability distribution with mean  $1/\lambda$ . The frequency distribution of service time revealed that it followed an exponential distribution and hence it was modeled as an exponential random variable.

- The arrivals followed Poisson distribution and occurred at an average rate of  $\lambda$  per hour.
- Each of the servers available provided service at an average rate of  $\mu$  per hour, and the actual service times followed an exponential distribution.
- Arrivals waited in a single queue and were served by the first available server.
- Finally,  $\lambda < \mu s$ .

The various queuing characteristics pertaining to M/M/s mode are:

$$U = \frac{\lambda}{s\mu} \quad (5)$$

$$\frac{1}{P_0} = \sum_{n=0}^{s-1} \frac{(\lambda / \mu)^n}{n!} + \frac{(\lambda / \mu)^s}{s!} \times \frac{s\mu}{s\mu - \lambda} \quad (6)$$

$$L_q = P_o \times \frac{(\lambda / \mu)^{s+1}}{(s-1)!} \times (s - \lambda / \mu)^{-2} \tag{7}$$

$$L = L_q + \frac{\lambda}{\mu} \tag{8}$$

$$W_q = \frac{L_q}{\lambda} \tag{9}$$

$$W = W_q + \frac{1}{\mu} \tag{10}$$

Using the aforementioned formulae, the average waiting time and the total average-time a customer spends in the system (in line and in service) were calculated.

Then, in the next two iterations, the number of servers was increased and again the values of those parameters were calculated and the comparison of three results was made. Thus, the calculations were done for up to 4 number of servers.

### 3. Calculations

Altogether, 49 customers visited the hair salon for the service within the considered time frame of 8 hours: 9 am to 5 pm, which is the normal working hour.

Table 1. Frequency distribution of service time

Service Duration (D, mins)	Frequency (f)	f.D
7	2	14
8	1	8
10	1	10
11	1	11
12	2	24
13	7	91
14	4	56
15	6	90
16	13	208
17	7	119
22	1	22
24	1	24
25	2	50
26	1	26
	$\Sigma f = 49$	$\Sigma f.D = 753$

Using (1), (2) and (3),

$$\lambda = \frac{49}{8} = 6.125 \text{ per hour}$$

$$D_{avg} = \frac{\sum f \cdot D}{\sum f} = \frac{753}{49} = 15.36 \text{ min}$$

$$\mu = \frac{60}{D} = \frac{60}{15.36} = 3.904 \text{ per hour}$$

For 2 servers, various queuing parameters are calculated using equations (5) to (10) as follows:

$$U = \frac{\lambda}{s\mu} = \frac{6.125}{2 \times 3.904} = 0.7844 \equiv 78.44\%$$

$$\frac{1}{P_0} = \sum_{n=0}^{\infty} \frac{(6.125/3.904)^n}{n!} + \frac{(6.125/3.904)^2}{2!} \times \frac{2 \times 3.904}{2 \times 3.904 - 6.125} = 8.2786$$

$$\therefore P_0 = \frac{1}{8.2786} = 0.12$$

$$L_q = 0.12 \times \frac{(6.125/3.904)^{2+1}}{(2-1)!} \times (2 - 6.125/3.904)^{-2} = 2.4935$$

$$L = 2.4935 + \frac{6.125}{3.904} = 4.0624$$

$$W_q = \frac{2.4935}{6.125} = 0.4071 \text{ hour} = 24.42 \text{ min}$$

$$W = 0.4071 + \frac{1}{3.904} = 0.6632 \text{ hour} = 39.79 \text{ min}$$

It is seen that, on average, a customer has to stay approximately 24 mins in the queue waiting for his turn, with his total time in system (the hair salon) being nearly 40 minutes.

Let's find out these two parameters again by increasing the number of servers (hairdressers) to 3 i.e.  $s = 3$ .

$$U = \frac{\lambda}{s\mu} = \frac{6.125}{3 \times 3.904} = 0.523 \equiv 52.3\%$$

$$\frac{1}{P_0} = \sum_{n=0}^2 \frac{(6.125/3.904)^n}{n!} + \frac{(6.125/3.904)^3}{3!} \times \frac{3 \times 3.904}{3 \times 3.904 - 6.125} = 5.1486$$

$$\therefore P_0 = \frac{1}{5.1486} = 0.194$$

$$L_q = 0.194 \times \frac{(6.125/3.904)^{3+1}}{(3-1)!} \times (3 - 6.125/3.904)^{-2} = 0.287$$

$$L = 0.287 + \frac{6.125}{3.904} = 1.856$$

$$W_q = \frac{0.287}{6.125} = 0.0468 \text{ hour} = 2.811 \text{ min}$$

$$W = 0.0468 + \frac{1}{3.904} = 0.303 \text{ hour} = 18.18 \text{ min}$$

Here, we can see that on increasing the number of servers to 3, the waiting time significantly drops down to 2.811 minutes from 24.42 minutes.

Now, let's repeat the process by increasing the number of servers to 4 to see the effect on the average times.

$$U = \frac{\lambda}{s\mu} = \frac{6.125}{4 \times 3.904} = 0.3922 \approx 39.22\%$$

$$\frac{1}{P_0} = \sum_{n=0}^3 \frac{(6.125/3.904)^n}{n!} + \frac{(6.125/3.904)^4}{4!} \times \frac{4 \times 3.904}{4 \times 3.904 - 6.125} = 4.85$$

$$\therefore P_0 = \frac{1}{4.85} = 0.206$$

$$L_q = 0.206 \times \frac{(6.125/3.904)^{4+1}}{(4-1)!} \times (4 - 6.125/3.904)^{-2} = 0.0552$$

$$L = 0.0552 + \frac{6.125}{3.904} = 1.624$$

$$W_q = \frac{0.0552}{6.125} = 9.012 \times 10^{-3} \text{ hour} = 0.54 \text{ min}$$

$$W = 9.012 \times 10^{-3} + \frac{1}{3.904} = 0.2651 \text{ hour} = 15.9 \text{ min}$$

It is seen that, on increasing the number of servers to 4, the average waiting time reduces to 0.54 minutes while the total time to be spent in the system reduces to 15.9 minutes.

It can be observed that increasing the number of servers further beyond 3 doesn't reduce the waiting time significantly in relative comparison with the number of servers. Just adding servers may result in increased expenses despite giving added benefit. Thus, it doesn't look attractive to increase the number of servers further.

#### 4. Summary

An analysis was made for the average time a customer had to spend waiting in a queue and on the system to get served in a hair salon. To find out the average arrival rate and the average service rate, the data were recorded for 8 hours from 9 am to 5 pm with an aid of the CCTV placed at the hair salon.

The arrival was observed to follow the random (Poisson) distribution whereas, the service time, the exponential distribution. The average arrival rate was calculated to be 6.125 per hour while average the service time to be 15.36 mins and the corresponding service rate to be 3.904 per hour.

The system was modeled as an M/M/s model as per the Markovian inter-arrival time and the Markovian service times. There were two servers (hairdressers) at the hair salon.

For 2 servers, the arrival rate of 6.125 per hour, and the service rate of 3.904 per hour, various queuing parameters were found out using the equations (5) to (10).

It was found that a customer had to wait, on average, about 24 mins in a queue for his turn and about 40 minutes in total at the hair salon to get his work done.

Then, the number of servers was increased to 3 and by the same procedure, the average waiting time and the total time to be spent, on average, in the system were determined. It was observed that the average waiting time reduced drastically and dropped down to 3 minutes (approx.) from 24 minutes (approx.).

Finally, when the number of servers was increased to 4, the average waiting time reduced to just 0.54 minutes with the total time in the system being 15.9 minutes. However, the decrement in waiting time doesn't look as convincing as it looked when it dropped from 24 minutes (approx.) to 3 minutes (approx.). Additionally, it may merely increase the cost of employing one more hairdresser.

Taking into consideration, the nearly-same flow rate of customers into the salon every working day, it is recommended that the owner employ 3 servers in total at the hair salon so that the customers need to wait only for about 3 minutes in a queue and get the whole task done in about 18 minutes.

#### 5. Conclusions

A steady-state analysis of the operation of a hair salon was carried out by modeling the system as an M/M/s model to identify the optimal waiting time for a customer and hence to make decisions on the total number of servers to be deployed with minimum expenses.

The optimal waiting time is found to be about 3 minutes for which 3 hairdressers (servers) are to be deployed.



Since, here, in Nepal, customers are not tempted to return due to just 3 minutes of waiting time, the presented solution seems optimum under the current situation of customer in-flow, labor expenses, and revenue generated from the customer service.

It is thus recommended that the owner employ 3 hair dressers so that a customer needs to spend about 18 minutes to get his hair done, where the waiting time in the queue is about 3 minutes.

Hence, in the areas with a similar pattern of customer flow rates (in the city), a total of 3 servers seems to be the best possible solution for optimum waiting time.

Nevertheless, it should be noted that under conditions where the labor expenses are too low in comparison to revenue generation, and high customer impatience (due to which customers may return after arrival), further analysis might be required to make the decision based on relative comparison between the benefit and the expenses/costs incurred.

A simple case was considered without taking into account, some scenarios: complex conditions where customers balk (not joining a queue upon arrival), or renege (leaving the queue before being served). Also, the arrival rate is considered uniform throughout the day.

### **Acknowledgements**

The work presented here is supported by the Department of Mechanical and Aerospace Engineering, IOE, Pulchowk Campus.

### **References**

- [1] P. Chen, Y. Lu, V. W. Zheng, X. Chen, and B. Yang, "KnowEdu: A System to Construct Knowledge Graph for Education," *IEEE*, vol. 6, pp. 31553–31563, 2018. [1] Ragsdale, C T.. Spreadsheet modeling & decision analysis: custom publication. Scarborough, Ont.: Thomson Nelson, 2008.
- [2] Bose, S K.. An introduction to queuing systems. New York: Kluwer Academic/Plenum Publishers, 2002.
- [3] Holt Jr, Charles A., and Roger Sherman. "Waiting-line auctions." *Journal of Political Economy* 90.2 (1982): 280-294.
- [4] Artalejo, Jesus R. "A queueing system with returning customers and waiting line." *Operations Research Letters* 17.4 (1995): 191-199.
- [5] Cernea, Sidonia Otilia, Mihaela Jaradat, and Mohammad Jaradat. "Characteristics of waiting line models-the indicators of the customer flow management systems efficiency." *Annales Universitatis Apulensis: Series Oeconomica* 12.2 (2010): 616.
- [6] Friedman, Hershey H., and Linda Weiser Friedman. "Reducing the wait in waiting-line systems: waiting line segmentation." *Business horizons* 40 (1997): 54-58.
- [7] Chowdhury, Mohammad Shyfur Rahman. "Queueing theory model used to solve the waiting line of a bank: a study on Islami Bank Bangladesh Limited, Chawkbazar Branch, Chittagong." *Asian Journal of Social Sciences & Humanities* 2.3 (2013): 468-478.
- [8] Levhari, David, and Israel Luski. "Duopoly pricing and waiting lines." *European Economic Review* 11.1 (1978): 17-35.
- [9] Meisling, Torben. "Discrete-time queueing theory." *Operations Research* 6.1 (1958): 96-105.
- [10] De Bruin, Arnoud M., et al. "Modeling the emergency cardiac in-patient flow: an application of queueing theory." *Health Care Management Science* 10.2 (2007): 125-137.
- [11] Kalashnikov, Vladimir V. *Mathematical methods in queueing theory*. Vol. 271. Springer Science & Business Media, 2013.

- [12] Dressin, S. A., and Edgar Reich. "Priority assignment on a waiting line." *Quarterly of applied mathematics* 15.2 (1957): 208-211.
- [13] Rao, S. Subba. "The relationship of work-in-process inventories, manufacturing lead times and waiting line analysis." *International Journal of Production Economics* 26.1-3 (1992): 221-227.
- [14] Kim, Bong-Sub, and Jae-Nam Lee. "A Study on The Characteristics of Aesthetic Salon Operation and Job Satisfaction by Important Aspects in Launching Aesthetic Salon Business." *Journal of the Korea Academia-Industrial cooperation Society* 19.2 (2018): 318-328.
- [15] Koh, Teow Hock, Zhenlin Yang, and Lijing Zhu. *The general dominance of lottery over waiting-line*. School of Economics and Social Sciences, Singapore Management University, 2002.
- [16] Luckhardt, Wesley E. "A Waiting Line Model: Determining the Number of Seats Needed in Dining Areas." *Journal of College & University Foodservice* 1.1 (1992): 25-37.

### Author's Profile



**Sambhav Kharel** (born December 15, 1996) is an undergraduate student in Mechanical Engineering at the Institute of Engineering (IOE), Pulchowk Campus. Kharel has opted for 'Operations Research and Management Science' as an elective during the final year of his graduation. He has been working enthusiastically in the field of Operations Research and Management Science. His email address is: 073bme631.sambhav@pcampus.edu.np



**Roshan Bhandari** (born September 7, 1997) is an undergraduate student in Mechanical Engineering at the Institute of Engineering, Pulchowk Campus. Bhandari is a committee member of Society of Mechanical Engineering Students (SOMES), Pulchowk Campus as an Event Manager. He has always been interested in the field of operational management. His email address is 073bme627.roshan@pcampus.edu.np.



**Satish K C** (born December 6, 1998) is an undergraduate student studying Mechanical Engineering at the Institute of Engineering, Pulchowk Campus, Nepal, taking Operation Research and Management Science as an elective. K C completed an internship at Varun Beverages Ltd, franchisee of PepsiCo. Satish will graduate in December 2020. His email address is 073bme637.satish@pcampus.edu.np.



**Aayush Bhattarai** (born January 18, 1990) is an Assistant Professor at the Department of Mechanical and Aerospace Engineering, Pulchowk Campus, Institute of Engineering, Tribhuvan University. Bhattarai is also a founding faculty of Aerospace Engineering for the first time in Nepal. He completed his M.Eng from the University of Technology Sydney, Australia, and Bachelor of Engineering in Aeronautics from Nanjing University of Aeronautics and Astronautics. His research interest mainly includes the field of Aviation, Operation Research, Operations Management, and Project Management. His email address is: aayush@pcampus.edu.np

**How to cite this paper:** Sambhav Kharel, Roshan Bhandari, Satish K C, Aayush Bhattarai. " A Steady-State Analysis of a Hair Salon as a Single-Queue, Multi-Server System to Optimize the Waiting Time in a Queue ", International Journal of Education and Management Engineering(IJEME), Vol.10, No.3, pp.22-32, 2020.DOI: 10.5815/ijeme.2020.03.03