Determining the Number of Optimum Servers in The XYZ Restaurant Queue System with Queuing Theory

Khawarita Siregar¹, Aulia Ishak², Fernando³

¹²³Department of Industrial Engineering, Faculty of Engineering, Universitas Sumatera Utara

E-mail: fernando.acc8@gmail.com khawarita@usu.ac.id

Abstract. Queue is a natural phenomenon that occurs when the demand for a service at a certain time exceeds the capacity of service at the same time. In this paper, the problem is solved using the queuing theory which is an analytical tool that is very helpful in solving queuing problems. This theory includes mathematical studies that produce important information needed in decision making with the help of forecasting various characteristics of the queue line. The queuing model in this restaurant queue system is (M / G / S). The current queuing system has System utility level (ρ) 52.23%, average time of customers in the queue (Wq) 0.4634 minute, The average time a customer in a system (W) 2.8074 minutes, Average Number of Visitors in the Queue (Lq) = 0.2104 ≈ 1 person, and Average Number of Visitors in the System (Ls) 1.2750 ≈ 1 person. The total waiting time on server 2 is below the maximum waiting time of 15 minutes and has the greatest system utility level, thus the optimum number of servers chosen is server 2 which is already optimum.

1. Introduction
Queue is a natural phenomenon that occurs when the demand for a service at a certain time exceeds the capacity of service at the same time. Queuing problems are often found in daily life both in industry and in trade, finance, social and others. Waiting for service situations will form a waiting line. These waiting lines are often called queues (queues), because service facilities (servers) are relatively expensive to fulfil service requests and are very limited. For example, in purchasing tickets in cinema, complaints are often encountered because the queue is too long where customers have to wait for a very long time where the payment counter can be opened but not opened and sometimes too many open counters that cause employees to idle. Extremely long queues and too long to get service are very annoying. On average, the waiting time depends on the average rate of service [1]. For this reason, the manager of the xyz restaurant wants to know the queuing system at the xyz restaurant is it good or needs to be repaired. The objective of this paper is to help the manager analyze and determine the number of servers that need to be opened in order to work more optimally. In this paper we analyze and optimize the queuing system using Queuing theory.

Queuing theory is an analytical tool that is very helpful in solving queuing problems. This theory includes mathematical studies that produce important information needed in decision making with the help of forecasting various characteristics of the queue line.

Application of queuing theory study conducted by Dimas Dwi Prayogo entitled "Analysis of queue systems and optimization of teller service At PT. Bank Sulutgo” discusses the analysis of the
application of the M / M / S model in the main branch of the North Sulawesi SulutGo queue system. The method used in this journal is M / M / S which is used like the average number of customers in the system, the average time customers spend in the system. It can be seen the Queue Discipline used by Bank SulutGo main branch implements the First Come First Serve (FCFS) system where customers who come first will be served first. The average number of customers in the queue occurred in the period of 12.00 - 13.00 where the average number of customers waiting in the queue was 1,385 people or = 1. But in the performance results table in the discussion of the average number of customers in the queue there was no waiting for the teller to be served immediately because it caused one teller to rest and the average standard of service level was 5 minutes then 60 minutes for 12 people were serviced. While the optimal number of tellers at the main SulutGo Bank branch is 5 tellers and it can be concluded that the performance of the queuing system at the main SulutGo Bank branch is optimal [2].

Another application of queuing theory study conducted by Seigha Gumus entitled “Application of Queuing Theory to a Fast Food Outfit: A Study of Blue Meadows Restaurant”. This study evaluated the queuing system in Blue Meadows restaurant to determine its operating characteristics and to reduce customers’ waiting time with queuing theory. In this study, they used M/M/S model for the queue, they evaluate the arrival rate 40 customers per hour, service rate was about 22 customers per hour per server. The servers that available is 2 with the utilization is highly above average at 0.909 that can be conclude very effective. The result of this study can be used as a reference to analyze the current system and improve next system [3].

Next application of queuing theory study conducted by Agustian Suseno entitled “Analysis of Queue System to Optimize Teller Services in BRI Bank Cibadak Branch, Sukabumi”. Due to the increase in the number of customers over time, responsiveness is needed to achieve customer satisfaction. Banking as a financial service in dealing with customers is closely related to the speed of service provided to customers, sometimes whether or not customers are served quickly is constrained by unpredictable queue times. The purpose of this study is to model the current queue system and compare it with the proposed improvement with the optimal queue size for the number of tellers at BRI Cibadak branch, Sukabumi Regency in order to improve service time effectiveness. When compared to the installation system for 5 tellers and 6 tellers, the results obtained are better for 6 teller installation systems with average server utilization of 67%, average number in the queue of 0.57 customers, average number in the system of 4.57 customers, average time in the queue is 0.01 hours and the average time in the system is 0.11 hours [4].

Another application of queuing theory is conducted on a bank too by Cut I.Setiawati entitled “Counting Teller Quantity For Better Queue In Financial Institution: Case of Bank Central Asia, Metro Indah Mall Branch Office-Bandung”. In this bank, which initially had 4 tellers, had the intention to improve the queuing system, because according to the company waiting customers are a disadvantage, by developing this system it is hoped that customer satisfaction can be achieved. After conducting this study showed that each teller could serve 10 persons per hour and teller busy the level is 97%. The average time spent by each customer in the system is 63 minutes and in the queue line is 57 minutes. This research results showed that there should be 7 tellers to get an optimal result, meaning that BCA Bank should provide three more tellers per day. Providing more tellers will indicate better benefits for the customer [5].

Based on the studies conducted earlier, Queuing Theory method can be used as a tool to analyze queuing system, improve queuing system, determine the optimal number of server. Therefore in this study the author will use Queuing theory to determine the optimum server in XYZ restaurant queuing system.

2. Method

The research approach used in this research is a quantitative approach as an approach in which is real and directly collected by observation. These data are arrival time data, service time data and exit time data which are observed with digital clock and recorded on the worksheet between arrival time data,
service level time and customer exit time from the queuing system for 8 hours and maximum customer waiting time by asking directly to the customer.

The queuing system consists of a queue line and a servant station. Customers who need service come from a source called the calling population entering the queue system from time to time. Customers come to the system and join in forming a queue. At a certain time, one of the members of the queue is chosen to be served. The selection is based on certain rules called disciplinary services. Services are provided to customers through a service mechanism and after they are served, the customer leaves the queuing system. Losses experienced by companies can be caused by many things, both errors caused by humans, machines, raw materials, work methods, and work environment. Therefore, we need a method that can support quality improvement with the aim of being able to avoid more losses and produce quality products [6]. There are four forms of service discipline commonly used in practice, namely:

- First In First Out (FIFO), which means that the first comes the first served.
- Last In First Out (LIFO), which means the one who arrives last, comes out first.
- Service In Random Order (SIRO), which means the summoning is based on random chance, it doesn't matter who arrives first.
- Priority Service (PS), which means service priority is given to those who have higher priority [7].

There are 4 basic structures of the queuing model that commonly occur in a queuing system:

- Single Channel Single Phase: indicates there is only one system entry service and there is only one service facility
- Single Channel Multi Phase: indicates that there is only one service entry line and there are two or more service facilities in series in the line
- Multi Channel Single Phase: indicates there are two or more system entry lines service and there is only one service facility in each line
- Multi Channel Multi Phase: indicates there are two or more system entry lines service and there are also two or more service facilities in series in each the path [8].

The data processing steps in this study are as follows:

- Testing the distribution of arrival frequency data, time between arrival data and service level time data with EasyFit Software.
- The queuing model is determined by Kendall Notation. Kendall notation is a standard notation which is a universal standard of combining arrival processes with services in the following format: (a / b / c): (d / e / f), a: arrival distribution (Arrival Distribution), b: distribution of service time or departure, c: number of services in parallel, with c = 1, 2, 3,… ∞, d: service discipline (eg, FCFS, LCFS, SIRO), e: the maximum number allowed in the system and f: number of customers who want to enter the system as a source [9].
- Queue system analysis which includes the average level of customer arrival (λ), average level of service(μ), level of system utilization (ρ), average number of visitors in the queue (Lq), average number of visitors in the system (Ls), average time of visitors in the queue(Wq), and average time of visitors in the system (W).
- The optimum number of servers is calculated.
3. Results and Discussion

The data collected will look for the type of distribution using easyfit software. First, arrival frequency data will be grouped at 6 minute intervals. Distribution test results on the arrival frequency can be seen in Table 1.

| #  | Distribution   | Kolmogorov Statistic | Smirnov Rank | Anderson Statistic | Darling Rank |
|----|----------------|----------------------|--------------|--------------------|--------------|
| 1  | Binomial       | 0.30624              | 2            | 6.8487             | 2            |
| 2  | D. Uniform     | 0.3125               | 3            | 34.758             | 4            |
| 3  | Geometric      | 0.42101              | 4            | 6.972              | 3            |
| 4  | Poisson        | 0.30002              | 1            | 5.9146             | 1            |
| 5  | Bernoulli      | No fit (data max > 1)|              |                    |              |
| 6  | Hypergeometric | No fit               |              |                    |              |
| 7  | Logarithmic    | No fit (data min < 1)|              |                    |              |
| 8  | Neg. Binomial  | No fit               |              |                    |              |

It can be seen from the table above that the closest distribution of Arrival Frequency is the Poisson distribution and the following graph that can be seen in Figure 1.

![Poisson graph on arrival frequency](image)

Figure 1. Poisson graph on arrival frequency

Second, the Service Level Time data is changed to minutes and sorted, then the distribution type will be tested. Distribution test results on the Service level time can be seen in Table 2.

| #  | Distribution | Kolmogorov Statistic | Smirnov Rank | Anderson Statistic | Darling Rank | Chi-Squared Statistic | Rank |
|----|--------------|----------------------|--------------|--------------------|--------------|-----------------------|------|
| 1  | Beta         | 0.04237              | 4            | 0.75638            | 6            | 7.5699                | 3    |
| 2  | Cauchy       | 0.13545              | 14           | 4.9538             | 13           | 18.793                | 12   |
|   | Distribution       | Chi-Squared  | df1  | df2  | Chi-Squared  | df1  | df2  |
|---|-------------------|--------------|------|------|--------------|------|------|
| 3 | Chi-Squared       | 0.2737       | 19   | 30.14| 19           | 106.9| 18   |
| 4 | Chi-Squared (2P)  | 0.15048      | 15   | 8.4213| 14           | 52.37| 15   |
| 5 | Exponential       | 0.22794      | 17   | 20.593| 16           | 70.868| 16   |
| 6 | Exponential (2P)  | 0.16045      | 16   | 9.7099| 15           | 51.078| 14   |
| 7 | Gamma             | 0.04173      | 3    | 0.51727| 1           | 5.8462| 2    |
| 8 | Gamma (3P)        | 0.04595      | 5    | 0.72924| 4           | 10.825| 5    |
| 9 | Laplace           | 0.11618      | 13   | 4.5051| 12           | 21.854| 13   |
|10 | Log-Logistic      | 0.07629      | 9    | 1.9187| 9            | 12.693| 7    |
|11 | Log-Logistic (3P) | 0.05876      | 7    | 0.89437| 7           | 15.171| 10   |
|12 | Logistic          | 0.08614      | 10   | 2.5619| 10           | 13.909| 8    |
|13 | Lognormal         | 0.06631      | 8    | 1.742 | 8            | 15.502| 11   |
|14 | Lognormal (3P)    | 0.05244      | 6    | 0.64126| 3           | 11.779| 6    |
|15 | Normal            | 0.09173      | 11   | 2.7079| 11           | 14.76 | 9    |
|16 | Student's t       | 0.69476      | 20   | 314.25| 20           | 1146.7| 19   |
|17 | Triangular        | 0.26077      | 18   | 27.184| 17           | 79.907| 17   |
|18 | Uniform           | 0.108        | 12   | 27.582| 18           | N/A  |      |
|19 | Weibull           | 0.0394       | 1    | 0.60213| 2           | 4.7282| 1    |
|20 | Weibull (3P)      | 0.04009      | 2    | 0.74367| 5           | 8.8238| 4    |

It can be seen from the table above that the closest distribution of Service Level time is the Weibull distribution and the following is the graph of Service Time level than can be seen on Figure 2.

**Figure 2.** Weibull graph on service time level
The restaurant queue system has a Poisson distribution frequency, and service levels are distributed with Weibull with 2 servers. The queuing model with Kendall notation is (M / G / 2) (FCFS / 4 / ∞), M: Frequency of arrivals with Poisson distribution, G: The level of service is distributed by Weibull, 2: Number of servers is 2, FCFS: First Come First Served are customers who come first served, 4: The maximum queue limit is 4 people and ∞: Infinite Population Width.

Average Customer Arrival Rate (λ) is based on the average number of customer arrivals at a fixed time interval. The number of time intervals every day is 6 minutes intervals for 8 hours. So the level of customer simulation results is:

\[ \lambda = \frac{N}{T} = \frac{218 \times 80}{8 \times 6} = 0.4542 \approx 1 \text{ person per minute} \]  
(1)

The average service time of each customer obtained in this observation is the division of the total service time divided by the amount of data for example the level of service, namely:

\[ \text{Average Service} = \frac{\text{Total Service Time}}{\text{Total Data}} = \frac{511.000}{218} = 2.344 \text{ minutes/person} \]  
(2)

Service Level (μ) = 1 / \text{Average Service} = \frac{1}{2.344} = 0.4266 \]  
(3)

In observation, there are 2 service facilities (servers). Based on the formula above, the system utility level of the observed object is:

\[ \rho = \frac{\lambda}{k\mu} = \frac{0.4542}{2 \times 0.4266} = 0.5323 \]  
(4)

%ρ = ρ x 100% = 0.5323 x 100% = 53.23%

Next will be calculated the average time a customer waits in the queue with the formula:

\[ W_q = \frac{\lambda^2 E(t)^{k-1}(k\lambda E(t))}{2(k-1)! (k\lambda E(t))^{1/2} \left( \frac{(k\lambda E(t))^{1/2}}{(k-1)! (k\lambda E(t))} \right)^{2/3}} \]  
(5)

With \( W_q \) is a service time in the queue line, \( \lambda \) is a average customer arrival rate, \( k \) is a number of servers and \( E(t) \) is a average service. Then, the average time of customers in the queue \( (W_q) \) is as follows:

\[ W_q = \frac{0.4542^2 (2.344)^2 (2.344)^{2-1}}{2(2-1)! (2-0.4542(2.344))} \left[ \frac{(0.4542(2.344))^{0}}{0!} + \frac{(0.4542(2.344))^{1}}{1!} + \frac{(0.4542(2.344))^{2}}{2(2-1)! (2-0.4542(2.344))} \right] \]

\[ W_q = 0.4634 \text{ minute} \]

The average time a customer in a system \( (W) \) is actually the time calculated from the customer entering the waiting line until the service process is complete and can be formulated as follows:

\[ W = W_q + \frac{1}{\mu} \]  
(6)

\[ W = 0.4634 + \frac{1}{0.4266} = 2.8074 \text{ menit} \]

Average Number of Visitors in the Queue \( (L_q) \) can be formulated as follows:

\[ L_q = \lambda (W_q) \]  
(7)

\[ L_q = 0.4542(0.4634) = 0.2104 \approx 1 \text{ person} \]

Average Number of Visitors in the System \( (L_s) \) can be formulated as follows:

\[ L_s = \lambda (W) \]  
(8)
Ls = 0.4542(2.8074) = 1,2750 person

Next is to calculate and comparing current amount of server with several amount of server such as 1 and 3 servers. The result of the calculation can be seen on Table 3.

| Server amount | ρ     | Lq    | Ls    | Wq    | W     | Maximum waiting time (minute) |
|---------------|-------|-------|-------|-------|-------|------------------------------|
| 1             | 106.46% | -8.7742 | -7.7096 | -19.3194 | -16.9754 | 15                           |
| 2             | 53.23%   | 0.2104  | 1.2750 | 0.4634 | 2.8074 | 15                           |
| 3             | 35.49%   | 0.0226  | 1.0871 | 0.0497 | 2.3937 | 15                           |

As can be seen in the table above, the time of the customer in the queue with 2 servers is 0.4634 minutes, this result is smaller than the customer maximum waiting time of 15 minutes and the system utility is 0.5323 which indicates that the system is good enough. Thus the optimum number of servers chosen is 2 servers.

4. Conclusion
Based on the results and discussion above, it can be concluded that Distribution type of Arrival Frequency is Poisson distribution and Distribution type of Service level time is Weibull distribution, and the The queuing model with Kendall notation is (M / G / 2) :( FCFS / 4 / ∞). The current queuing system has System utility level (ρ) 52.23%, average time of customers in the queue (Wq) 0.4634 minute, The average time a customer in a system (W) 2.8074 minutes, Average Number of Visitors in the Queue (Lq) = 0.2104 ≈ 1 person, and Average Number of Visitors in the System (Ls) 1.2750 ≈ 1 person. The total waiting time on 2 server is below the maximum waiting time of 15 minutes and has the greatest system utility level, thus the optimum number of servers chosen is 2 servers.

Acknowledgment
The author would like to thank University of Sumatera Utara who has Provided financial support to publish this paper. The author also thanks all who have participated and contributed in completing this paper.

Reference
[1] Kakiay T J 2004 Dasar Teori Antrian Untuk Kehidupan Nyata (Yogyakarta: Penerbit Andi)
[2] Prayogo D 2017 Analysis of Queue System and Optimization of Teller Service at PT Bank Sulutgo Jurnal EMBA 5 pp 928-934
[3] Gumus S 2017 Application of Queuing Theory to a Fast Food Outfit: A Study of Blue Meadows Restaurant Independent Journal of Management & Production 8 pp 441-458
[4] Suseno A 2020 Analysis of Queue System to Optimize Teller Services in BRI Bank Cibadak Branch, Sukabumi Journal of Industrial Engineering and Management Systems 13 pp 41-46
[5] Setiawati Cut 2017 Counting Teller Quantity For Better Queue In Financial Institution: Case of Bank Central Asia, Metro Indah Mall Branch Office-Bandung Jurnal Bisnis & Manajemen XVIII pp 65-72
[6] Sukaria Simulingga 201 Pengantar Teknik Industri (Medan: USU Press)
[7] Arwindy F 2014 Analisis Dan Simulasi Sistem Antrian Pada Bank ABC Saintia Matematika 2 p 2
[8] Subagyo P 1984 Dasar-Dasar Operation Research (Yogyakarta : BPFE)
[9] Ratnakusuma A 2015 Analisis Sistem Antrian Pesawat Terbang Di Bandara Internasional Ahmad Yani Semarang Jurnal Gaussian 4 pp 725-733
[10] Diah Puspitasari 2005 Aplikasi Sistem Antrian dengan Saluran Tunggal pada Unit Pelaksana Teknis (UPT) Perpustakaan Universitas Negeri Semarang