An analysis on Kendall Lee queueing system with non-preemptive priority at BRI Ahmad Dahlan Yogyakarta

S D Ayuningtyas and N Binatari

Department of Mathematics Education, Faculty of Mathematics and Natural Science, Yogyakarta State University, Jl Colombo No.1 Depok Karangmalang Sleman Yogyakarta Indonesia

E-mail: nikenasih@uny.ac.id

Abstract. This research was aimed to determine the Kendall Lee model, measure the Performances and analyze the optimal condition of queue at tellers line at BRI Unit K.H Ahmad Dahlan Yogyakarta. First, the model was constructed based on Kendall Lee in which the parameter values were obtained from the fitness test of the data distribution. Next, the Performances measurements were counted using Little's Formula while the optimal condition was hold based on the total cost. Since the addition of the number of servers will influence the total cost, the optimal condition is reached for a number of servers with lowest total cost and satisfying steady state condition. The result shows that the queue system in teller counter at BRI Ahmad Dahlan Yogyakarta uses M/M/2 model with non-preemptive service distribution. The optimum number of the server due to its total cost was found to be two servers. By these, it could be concluded that the system has been using optimal condition.

1. Introduction

In our everyday life, the phenomenon of the queue is very common. A queue is a part of the circumstances that occur in a series of random operational activities in a service facility [1]. The queue occurs because the need for a service exceeds the capacity available for the service. Therefore, to increase the system’s performance, the service providers provide multiple servers. Moreover, some system differs the customers into different priorities such as in customer service centers, airport security checkpoints, hospital emergency rooms, cloud computing systems or processor management in certain computer operating systems [2].

In a priority queue, each customers has a relative priority to be served [3]. Based on its priority level, there are two types of priority service discipline. The first one is the non-preemptive priority. If during the service time, one with higher priority customer arrives, it must wait until the item of lower priority has been served [4]. The second one is the preemptive priority. If during the service time, one with higher priority level arrives, it immediately displaces the lower priority customer. In this case, priority service discipline brings new problem on the queueing system.

The queue theory is one of the branches of applied mathematics that can be used to solve queuing problems. According to Heizer and Render, this has been very usefull for a service or manufacturing company [5]. One model of queue system was introduced by Kendall, which later was developed by Lee. The line model of Kendall Lee is a model line that is used to specify the characteristics of a queue includes the distribution of arrival and departure of customers, the number of servers, service
discipline, the capacity of the system and the number of customers who want to enter the queue system as the source. The used Kendall Lee line model is denoted by \((a/b/c) : (d/e/f)\) [6], where if the system capacity and calling source are both infinite then it is simplified by \(a/b/c/d\).

A research study on the system with priority service discipline has been done by numerous researchers. For the single server, it has been discussed in [3]. Specifically, a system with two priority classes has been solved in [4] while the waiting time for markovian multi-server with non-preemptive priority \(M/M/c\) was discussed in [7]. Research on the system of service priority queue with non-preemptive multi-server was done by Gail, Hanlter, and Taylor [8]. From the research, they have derived the Performances measurements of the system and the waiting time for each priority class. A research conducted by Kao and Wilson [9] also discusses the customer waiting time for the system of non-preemptive priority queue multiple servers with two priority classes. Another study that discusses queuing system with a priority of service was done by Ni'amah and Sugito [10].

A queue with non-preemptive priority was implemented by a service provider namely BRI KH Ahmad Dahlan Yogyakarta at teller’s line. The abundance of customers from the customer service’s line and surveyor’s line to teller’s line has a higher priority to be served than the direct customers. By these, they found their direct customer service’s delays are lining up to get service. The service delays occur whenever there is an abundance of customer service. Therefore, an analysis of this queueing system should be done.

In this research, it will be discussed the Kendall Lee model with non-preemptive priority occurred in BRI Unit KH Ahmad Dahlan Yogyakarta. The purpose of this study was to analyze the model line for Kendall Lee, to measure the Performances of the service system at teller line, and to find the number of server which optimize the total cost.

2. Method
The following will explain the methods and research design, location and time of the study, data collection techniques and design, and data analysis techniques.

2.1. Method and research design
The method used in this research was descriptive research method. Descriptive research method aims to describe the character of a variable, group or social phenomena that occur in society [11]. The design used in this research was the case and field research design.

2.2. Data collection technique
The research was conducted at PT Bank Rakyat Indonesia Unit KH Ahmad Dahlan Yogyakarta which is located at KH Ahmad Dahlan st No. 08 Yogyakarta. The process of retrieval data was obtained in the customer's waiting room, figure 1. The data was collected using the primary data collection technique in two ways, by using interview and field observation. The data are the interarrival time of abundance customers from customer service, the average rate of customer arrival of the queue directly to the teller and the service time.

![Figure 1. Queue system.](image-url)
2.3. Data analysis technique
Customer data analysis techniques in teller line BRI Unit KH Ahmad Dahlan in Yogyakarta is as follows:

1. The primary data for the number of direct customer arrival, abundance customer arrival time and the service rate are grouped per 15 minutes and obtained for 4 hours of research per days.
2. Poisson distribution test for the data was conducted using the Kolmogorov-Smirnov test. According to Siegel [12], the calculation steps for Kolmogorov-Smirnov test are (a) define the theoretical cumulative function, i.e. the cumulative distribution under $H_0$, (b) observe the values by pairing each interval $S_{(N)}(X)$, the cumulative frequency distribution which is obtained from a random sample of $N$ observations, with intervals $F_{(0)}(X)$, the cumulative distribution function under the $H_0$ assumption, and then subtract $F_{(0)}(X)$ with $S_{(N)}(X)$ for each level. (c) count the test statistic, 

$$D = \max \left| F_{(0)}(X) - S_{(N)}(X) \right|.$$ 

(d) determine the $D$ table value which is determined by referring to the critical values from the one sample Kolmogorov-Smirnov table.
3. The procedure to conduct a one sample Kolmogorov-Smirnov test was:
   - Null hypothesis $H_0$: the data follows the Poisson distribution,
   - Alternative hypothesis $H_1$: the data do not follow the Poisson distribution.
   - The level of significance $\alpha$,
   - Critical value: the null hypothesis, $H_0$, is rejected if the test statistic $D$ is greater or equals than the critical value obtained from the table, $D \geq D_{table}$.
4. Determine the queue model of the customers queueing system in teller line at Bank Rakyat Indonesia (Persero) Branch Office Unit KH Ahmad Dahlan Yogyakarta. According to Taha [6], Kendall Lee's Queue Model is denoted as follows:

$$(a/b/c):(d/e/f).$$

where
- $a$: Distribution of the arrival rate
- $b$: Distribution of services or departures time rate
- $c$: Number of server in parallel, $c = 1, 2, 3, ...$
- $d$: Service disciplines, such as FCFS, LCFS, SIRO
- $e$: The maximum capacity of the system
- $f$: The calling source
5. The traffic intensity for the $i$-th priority customer is $\rho_i = \frac{\lambda_i}{c \mu}$, where $\mu$ is the average rate of customer services, $\lambda_i$ is the average of customer arrival rate on $i$-th priority, and $c$ is the number of tellers to serve. Examine the steady-state solution, $\rho = \rho_1 + \rho_2$. It is exist if the value is less than one. In other words, the system has reached a stable condition.
6. Calculate the Performances of the queueing system. The measures of Performances are the expected value of customer waiting time on all priorities ($W_q$) and the expected number of i-th customers in the queue system ($W_q^{(i)}$). According to Gross et al [13], $W_q$ and $W_q^{(i)}$ can be calculated successively using equation (1) and equation (2).

$$W_q = \sum_{i=1}^{2} \frac{\lambda_i}{\lambda}W_q^{(i)} \quad (1)$$
and

$$W_q^{(i)} = W_{q0}^{(i)} + \frac{1}{c \mu},$$

where

$$W_q^{(i)} = \frac{c!(1-\rho)(c\mu) \sum_{n=0}^{c-1} \frac{(c\rho)^{n-1}}{n!} + c\mu}{(1-\sigma_{i-1})(1-\sigma_i)}$$

and $\sigma_i = \sum_{k=0}^{i} \rho_k$, $\sigma_0 = 0$ \( (3) \)

6. Optimize the queue system by considering its total cost. The cost analysis has taken a significant place in the queueing theory \[14\]. The total cost was defined as the sum total of the total cost of services per unit time, and the cost of the waiting customers per unit time \[5\]. In mathematical symbol, the cost model can be counted using the following formula:

$$ETC(c) = ESC(c) + EWC(c)$$

where

$c$ : the number of servers

$ETC(c)$ : the total costs for $c$ servers per unit time.

$ESC(c)$ : the expected service cost per unit time for $c$ servers.

$EWC(c)$ : the expected waiting cost per unit time for $c$ servers.

According to Gross et al \[13\], to calculate the value of the expectations of many customer types are in the queue system, it can be using the equation 5:

$$L_q^{(i)} = \frac{\lambda c!(1-\rho)(c\mu) \sum_{n=0}^{c-1} \frac{(c\rho)^{n-1}}{n!} + c\mu}{(1-\sigma_{i-1})(1-\sigma_i)}$$

\( (5) \)

3. Research and Discussion

The queueing in teller line at BRI Unit KH Ahmad Dahlan follows priority discipline. The abundance customer from the customer service line and the surveyor line to the teller line has higher priority to be served. Therefore, the direct customer in the teller line should wait until the abundance customer are completely serviced. In otherwise, if the arrival abundance customer occurs when all the teller are serving the customers, then the abundance customer should wait until the service is completely done. This type of priority discipline is called a nonpreemptive priority service discipline.

The study of this research deals only for teller service. Once customer completely served from the teller line, it is assumed that the customer left the queueing system. Therefore, the queue follows single phase multi channel model. The queue and service flow model can be seen from figure 2.

![Figure 2. Queue model.](image-url)
3.1. Distribution Match Test
Distribution suitability test was used to test the data directly teller line customer arrival, arrival overflow customers, and customer service is the Kolmogorov-Smirnov test. It was obtained that the direct customer arrival per 24 minutes was Poisson distributed with parameter $\lambda_1 = 8$ the abundance customers arrival was following Poisson distribution with $\lambda_2 = 1$ and the service was exponentially distributed, $\mu = 9$.

3.2. Kendall Lee’s Queue Model
Teller line systems in PT Bank Rakyat Indonesia (Persero) Branch Office Unit KH Ahmad Dahlan Yogyakarta has a model line of Kendall Lee $(M/M/2):(NPD/\infty/\infty)$. This model means arrivals and customer service Poisson distributed with 2 servers. Service discipline on the model $(M/M/2):(NPD/\infty/\infty)$ contains non-preemptive priority rule with unlimited system capacity and calling source.

3.3. Steady State Check
The steady state condition is reach if the traffic intensity is less than one. 

$$\rho = \frac{\lambda_1 + \lambda_2}{2 \cdot \mu} = \frac{1}{2} < 1$$

From this result, the queuing system with two server has reached steady state condition.

3.4. Performances measurements
The system characteristics could be recognised by looking at the expected rate of waiting time of the customer in queue line for both priority, $W_q$, and the expected rate of time spent by $i$-th type of customer on the system, $W_s^{(i)}$. By substituting the parameter’s values into equation (1) and equation (2), the performances measurements from the system are as presented in table 1.

| Table 1. Performance’s measurements |
|-----------------------------------|
| Measurements | Value   |
| $W_q^{(1)}$ (minute) | 1.4084 |
| $W_q^{(2)}$ (minute) | 1.9746 |
| $W_q$ (minute) | 1.0614 |
| $W_s^{(1)}$ (minute) | 2.1584 |
| $W_s^{(2)}$ (minute) | 2.7246 |

3.5. Cost Analysis Queue
In order to provide a better service, management could perform three ways, add the number of servers, decrease the service time or reduce the system capacity. This research only considered with the number of servers by using the cost model. The cost model can be used to evaluate and determine the optimum number of servers in the queue system [15]. The operating costs to be considered in this
research were the cost of waiting for the customer, the salary and the logistics costs for adding a number of tellers. Further information, teller in PT Bank Rakyat Indonesia (Persero) Branch Office Unit KH Ahmad Dahlan Yogyakarta work for seven hours each day and five days in a week.

There was an extra cost for servicing if there is an additional teller. The expected service cost per unit time is equals to the number of server which is multiplied with the service cost for one server. In other word, the formula for the expected service cost can be written as equation (6).

\[ ESC(c) = c \times C_i \]  

(6)

Now, we considered to the value of the service cost for each server. To serve the customer, the management should pay for the salary and the logistic cost. The salary for a teller per month is Rp 3,500,000 and the logistic cost to buy a computer, desk, chairs, and printer are Rp 4,000,000, Rp 1,100,000, Rp 700,000 and Rp 1,700,000, respectively. Assumed that the life period of the logistic was five years, then the service cost per minute is given in table 2.

| Table 2. Cost of services |
|----------------------------|
| **Cost Type** | **Unit** | **Cost per Month (Rp)** | **Cost per Day (Rp)** | **Cost per Hour (Rp)** | **Cost per Minute (Rp)** |
| Salary | Person | 3,500,000 | 175,000 | 25,000 | 417 |
| Computer | Set | 67,000 | 3,350 | 479 | 8 |
| Table | Unit | 19,000 | 950 | 136 | 3 |
| Chairs | Unit | 12,000 | 600 | 86 | 2 |
| Printer | Unit | 29,000 | 1,450 | 208 | 4 |
| **Total** | | 3,627,000 | 181,350 | 25909 | 434 |

From the table, it can be concluded that the service cost per minute for one server is Rp 434.

Next, the expected waiting cost is considered to the customer’s disadvantage for being in the system. Suppose that in order to be served, the customer should leave from work. In other words, the cost of waiting equals the cost of the customer if it is not going to work. Therefore, the waiting costs for customers are related to the per capita income of Indonesia. According to the Central Bureau of Statistics per capita income of Indonesia in 2016 was Rp 47.96 million. Therefore, the cost of waiting customer per minute, \( C_2 \), was Rp 476. The expected waiting cost per unit time

\[ EWC(c) = L_s \times C_2. \]  

(7)

Table 3 presents the three scenario of number of tellers, which are one, two and three, and the traffic intensity, \( \rho \).

| Table 3. The expected number of customer |
|----------------------------------------|
| **Number of Teller (c)** | **\( \lambda \)** | **\( \mu \)** | **\( \rho = \frac{\lambda}{c\mu} \)** | **Steady State** | **\( L_s \) (customer)** |
|-------------------------------|-----------------|-----------------|-----------------------------|-----------------|-----------------|
| 1                             | 77              | 77              | 1                           | No              | -               |
|                               | 240             | 240             |                             |                 |                 |
| 2                             | 77              | 77              | 1                           | Yes             | 1,4203          |
|                               | 240             | 240             | 2                           |                 |                 |
| 3                             | 77              | 77              | 1                           | Yes             | 1,0988          |
|                               | 240             | 240             | 3                           |                 |                 |

Based on the table 3, for one teller, the system will not reach a stable condition. It means that the teller will have no time for idle or rest. Since the optimal condition is reach if the traffic intensity is less than 1, a one teller queue system is not an optimal condition. Next, the system’s stability was found to be in minimum two tellers. The greater number of servers, the less expected number of
customers in the system. Therefore, we determined the optimum number of servers regarding to the number satisfied steady state condition but also minimize the total cost. Table 4 shows the total needed cost if there is an additional teller.

| Table 4 Total cost per minute |
|-----------------------------|
| Total | Service Cost | Total Service Cost | Waiting Cost for Customers | Number of Customers in the System | Total Waiting Cost | Total Cost |
| c | $C_1$ | $c \times C_1$ | $C_2$ | $L_s$ | $L_s \times C_2$ | $(c \times C_1) + (L_s \times C_2)$ |
| 2 | Rp 434 | Rp 868 | Rp 476 | 1.4203 | Rp 677 | Rp 1.545 |
| 3 | Rp 434 | Rp 1,302 | Rp 476 | 1.0988 | Rp 524 | Rp 1.826 |

Based on the calculation of the total cost per minute in table 4, the optimal number of tellers was two tellers because it requires the lowest total cost, accounted to Rp 1,545. Since the number of tellers at the moment was 2 systems, we can conclude that the queue system was already under optimal conditions.

4. Conclusion
Based on the results, the queue system for teller service at PT Bank Rakyat Indonesia (Persero) Branch Office Unit KH Ahmad Dahlan in Yogyakarta follows multiple channels single-phase line models Kendall Lee ($M/M/2/NPD$). It means that the arrivals and service are Poisson distributed, two servers are available and the applied service’s discipline is a non-preemptive priority. In addition, the capacity of queuing systems and calling sources is unlimited. The size of the Performances of the queue system in the company includes the average expected value of the customer's waiting time during 1.0614 minutes with an average value of the expected timing of the overflow customers and clients direct line tellers are in the queue until the system had been served consecutively for 2.1584 min and 2.7246 minutes. Teller queue system in the enterprise can achieve optimal conditions currently served by two tellers. Therefore, the queue system is already under optimal conditions.

References
[1] Kakiay T J 2004 Dasar teori antrian untuk kehidupan nyata (Yogyakarta: Andi)
[2] Brandwajn A and Begin T 2017 Multi-server preemptive priority queue with general arrivals and service times Performance Evaluation (Amsterdam: Elsevier)
[3] Miller R G 1960 Priority Queues. Ann. Math. Statist. 31 86
[4] Chang W 1965 Queuing with nonpreemptive and preemptive-resume priorities. Operations Research 13 1020-22
[5] Heizer J and Render B 2011 Operations Management: 10 (New Jersey: Pearson).
[6] Taha H A 2007 Operation Research and Introduction (Upper Saddle River: Pearson Education, Inc)
[7] Kella O and Yechiali U 1985 Waiting times in the non-preemptive priority M/M/C queue. Stochastic Models, 1 257-62
[8] Gail H R, Hantler S L et al 1988 Analysis of a non-preemptive priority multiserver queue Advances in Applied Probability 20 852-79
[9] Kao P C and Wilson S D 1999 Analysis of non-preemptive priority queues with multiple servers and two priority classes. European Journal of Operational Research 118 181-93
[10] Ni'amah D and Sugito 2011 Queuing system with priority service Proceedings of the National Seminar on Statistics (Semarang: UNDIP)
[11] Martono N 2011 Quantitative research methods (Jakarta: Rajawali Pers).
[12] Siegel S 1956 Nonparametric Statistics for the Behavioral Sciences (New York: Mc Graw-Hill)
[13] Gross D 2008 Fundamentals of Queuing Theory: 4. (New Jersey & Sons, Inc: John Wiley).
[14] Mishra S S, Yadav D K 2010 Computational approach to cost and profit analysis of clocked queueing networks *Contemporary Engineering Sciences* 3 365-70

[15] A Muhajir, N Binatari 2017 Queueing System Analysis of Multi Server Model at XYZ Insurance Company in Tasikmalaya City *AIP Conference Proceedings* 1868