Perner's Contacts ***17****(2), 2022*
DOI: 10.46585/pc.2022.2.2400

OPTIMIZATION OF TOLL SERVICES USING QUEUING THEORY IN THE CASE OF ETHIOPIA

Abate Sewagegn[1,*]

Abstract *The toll road service is planned and built to fund road construction and management of traffic operations. Although the toll service has several benefits, it also has challenges by creating a queue in the event of high traffic. Queues at a toll service in addition to pollution and waiting time costs also expose drivers to road traffic accidents and affect the economy of the country. Queuing problems at highway gates and exits due to toll service have become the main concern for transport managers and planners because of the randomness of inter-arrival and service time. The critical point in the management of toll road services is finding the optimal number of servers. This paper aims to analyze the queue feature to optimize the toll service using analytical and simulation models by considering the Addis-Adama expressway toll service in Ethiopia. The paper used a mathematical model and a simulation model using the CPN tools to investigate the queuing parameters. Then the model results helped us to get the optimum output by using economic analysis and identifying the minimum waiting time and operating cost without expense. The performance of the toll service highly depends on the number of servers, the number of vehicles in the system and queue, waiting time in the queue, service time and inter-arrival time.*

Keywords *Queuing theory, Simulation, Service optimization, Toll, Model, CPN tools*

1 INTRODUCTION

In recent years, there has been an increasing interest in toll roads, i.e., roads that are planned and built with the intention of collecting income from road users to finance road construction and operation (R. Krol, 2018; Smitha, 2019; Board, 2006) and used as demand traffic management (Mckinnon, Knodler and Christofa, 2014). The Ethiopian transport system is one of the lowest in the world and insufficient for mobility. The problem is related to the weak institutional capacity like finance for construction and maintenance of road transport.

In developing countries like Ethiopia, the road infrastructure development source of finance is the public sector as a result the road infrastructure is underdeveloped. To reduce the problem the Ethiopian government starts building an expressway in 2014 to avert the increasing traffic demand with toll service. The first six-lane expressway toll road is constructed from Addis-Adama which is 84 Km and it's a vital route for import and export logistics. Because of this reason the expressway provides service to more than 20000 vehicles per day approximately a vehicle per 5 seconds. So the toll system helps the government to collect fees and provide better mobility(Abera, 2019). The traffic demand for this expressway increase from time to time as the vehicle population is increasing. However, the toll service point creates traffic congestion, especially at pick hours. Research by Abubakar et al., 2016 and H. Xiao and G. Zhang, (2010) indicate that even if the provision of a toll service has benefits, it also creates

[1] VSB – Technical University of Ostrava, Faculty of Mechanical Engineering, Institute of Transport, 17. listopadu 2172/15, 708 00 Ostrava, Czech Republic
[*] corresponding author, phone: +420 773 197915, e-mail: abategetaw4@gmail.com

problems at the entrance and exit of the road by increasing traffic and creating queues due to the imbalance between the arrival rate and the service. For this reason, road users waste their time at toll places. Compared to other factors such as cost, distance, quality, and speed, queueing time is the most critical element that affects the provision of efficient and effective service (Abubakar et al., 2016). When there is a long queue at a toll service, in addition to pollution, fuel cost, and waiting time cost, it exposes drivers to accidents and affects the country's economy (Duhan et al., 2014).

To minimize such queue problems, the analysis of the performance of toll services using queuing theory plays a critical role. Queuing theory is the study and modeling of waiting time or sequences of lines (H. Xiao and G. Zhang, 2010; Prasad et al., 2018). The queuing model is a useful mathematical technique for resolving a variety of transport queuing problems in any system, as it focuses on the description of traffic situations using a variety of mathematical models and formulas (Varghese and Chandran, 2021). Studies such as Molla, (2017); Wang (2017) apply queuing theory to analyze the mean number of customers served at one time, the mean waiting time in the queue, the average time in the system, the expected queue length, and the probability that the system is in a definite state, such as idle or busy.

Previous studies have used queuing models to analyze toll services; for example, Wang, (2017) tried to evaluate the performance of toll services using M/M/1 queuing theory by applying different service systems. Similarly, the authors Boahen, Adams and Salifu, (2013) apply queue theory to analyze the toll service in Accra, and their results showed that manual payment increases queue length in contrast to automatic service. Research by (Aksoy, Celikoglu and Gedizlioglu, 2014) indicates that a decreased number of toll services has an advantage by increasing the capacity of the bottleneck. In addition to this, the researchers (Yan and Lam, 1996) attempted to study toll patterns to reduce traffic congestion at peak hours, and traffic congestion on a general road network with both queues and congestion by combining traffic assignment and road prices. Duhan et al., (2014) use the queuing theory based on parameters to analyze the current situation of traffic congestion on highway services. However, these studies focus on analyzing toll services using parameters or simple models that do not show the dynamic features of the queue and do not address the issue of how many servers are required to optimize a toll service. In addition to a few studies conducted in the area of toll service, these research studies did not combine the mathematical model with the simulation model that allows iterating and cross-checking the validity of the model with the real system which is dynamic by its nature. According to Nico et al., (no date) combining a mathematical model with simulation helps us to test/ investigate different parameters without cost. A hybrid approach analysis of the toll service by combing simulation and queuing theory provides powerful results by using the conceptual framework of toll service from queuing theory and simulation used to evaluate and compare different parameters by iteration.

The main aim of this paper is to optimize the performance of toll services by analyzing steady-state solutions for the infinite capacity of a multi-server queuing system with identical capability. Therefore, by optimizing the performance of the toll service, it is possible to solve traffic congestion and reduce the number of traffic accidents by evaluating the variables of the queuing system using the mathematical analysis and simulation model with the CPN tool. Moreover, this study provides a springboard for the construction of an effective feature toll road that optimizes the toll service.

The paper is structured as follows: Section 2 describes the methods and model; in Section 3, mathematical equations are first developed using steady-state probabilities, and next, we examine the toll service's performance. In Section 4, we describe the simulation model developed using the CPN tool. First, we develop a conceptual model of the real-world system and then simulate the performance indicators using different parameters. In Section 5, results and discussions are presented and finally, conclusions are presented in Section 6.

2 METHODOLOGY AND MODEL

2.1 Description of the model

As we stated in the Introduction, toll service is designed to finance road construction and manage traffic congestion, but it has its drawbacks. For example, drivers are exposed to waiting in the queue to get a toll service and congestion after the service due to merging problems (Cherng, Lewis and Pai, 2005). When the number of toll booths increases, in addition to operating costs, the merging problems also increase.

The problem is multidimensional unless the design of toll service systems is systematically studied to employ the number of toll booths. For example, there is a driver cost related to waiting time and fuel cost, an environmental cost related to CO2 and noise pollution, and an operating cost. As traffic movement is sensitive, one change may affect the whole traffic system. Therefore, to minimize the problem of toll service and maximize benefits, an optimal number of toll servers is vital. As it has been established by Astutik and Dewanti, (2020) the toll queue model is a systematic representation and analysis of toll services to manage waiting time and service delivery.

The Addis-Adama expressway is designed with a multi-server that has a single queue as the vehicle in the queue from one lane can switch to the other lane, so to analyze the problem; we consider the M/M/c queuing model. The first **M** stands for the inter-arrival time distribution, and the second for the service time distribution. They are both exponential and thus memory less (Markovian). The letter **c** represents the number of servers (counters).

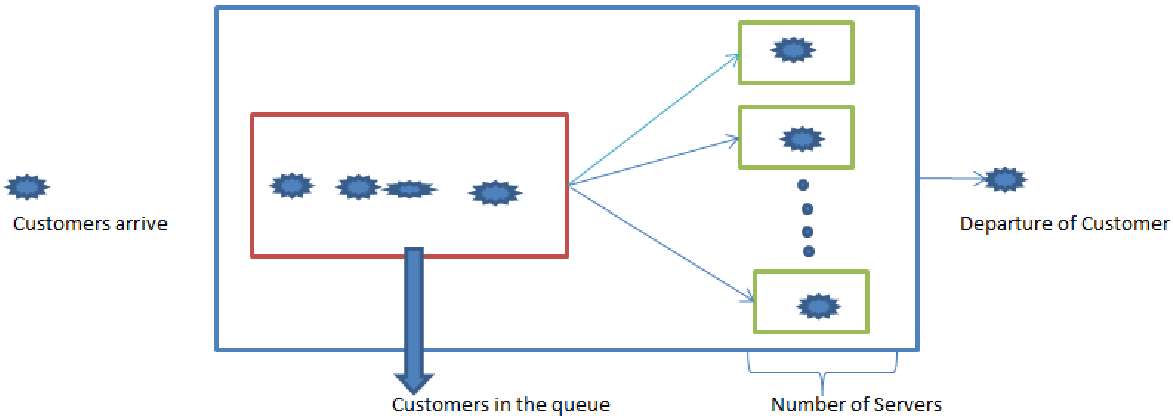

Fig. 1 M/M/*c* toll queuing model; source: author

The arrival of vehicles at the toll station and the service time taken by a vehicle are randomly distributed. Therefore, the vehicle's inter-arrival follows an exponential distribution with a mean inter-arrival time of $1/\lambda$ and the service time is exponentially distributed, expressed in mean service time $1/\mu$. The paper by Morgan, Banks and Carson, (1984) indicates the Poisson arrival process is an important one for an infinite population model inter-arrival time of a successive customer. In addition to this, the toll service uses a manual service system by a counter or cashier. The service is provided for the customer based on their arrival without jumping situation i.e. first in, first out. We are going to consider various queue parameters, such as the average number of users in the queue, the number of arrival per minute, the waiting time of users in the queue and the system. Taking such data into account, a mathematical and simulation model is developed. Finally, we select seven counters that optimize the toll service.

General assumptions and notations used in the model:

- The counter or the server works without failure or vacation.
- Each server has an identical capacity to serve the customer.
- The user stays in the system until they receive the intended service.
- n = mean number of vehicles/ unit time.
- c = the number of parallel servers.
- p_n = the probability of n vehicles in the system.
- μ = the mean service rate, which is expressed as vehicles/ unit time.
- λ = the mean arrival rate of the Poisson input process, which is expressed in vehicle/unit time.
- ES= the mean number of vehicles in service.
- EL = the mean number of vehicles in the system.
- ELq = the mean number of vehicles in the queue.

- EW = the mean waiting time for vehicles in the system.
- EWq = the mean waiting time of vehicles in the queue.

2.2 Characteristics of the Toll Queue Model

The queuing model is characterized by the arrival of the customer, service pattern, the number of servers, system capacity, and queue discipline.

1. Arrival pattern: - it is the distribution of the inter-arrival time of the customers. The mean value of the time between arrivals is vital and its reciprocal is the arrival rate of the customer, denoted by 'λ'.
2. Service pattern: - it is the time taken to serve the customer. The inverse of the mean service time is the service rate represented by 'μ'.
3. The number of servers: - the numbers of servers that provide service either in parallel or series, which can be single or multi-server.
4. The capacity of the system: - the system can be finite or infinite capacity.
5. Queue discipline: - There are different queue disciplines such as First-In-First-Out (FIFO), Last-In-First-Out (LIFO), service in random order (SIRO), and priorities. Among these disciplines, the model used First-In-First-Out (FIFO).

The main goal of queue studies is to examine the performance of service systems using variables such as inter-arrival rate, service rate, number of servers, service pattern, busy and idle servers, server utilization, service discipline, number of customers in a line and the system, waiting time in line, and time spent in the system, among others, to determine the system's efficiency and capacity (Astutik and Dewanti, 2020; Morgan, Banks and Carson,1984).

This paper used the steady-state behavior of the system rather than transient because the state of the system is categorized as a transient state when its operating characteristics, arrival, waiting time, and service time of the customer depend on time while in a steady state they are independent on time. In the explosive state, when the arrival rate of the system is more than its service rate and the queue increases with time and leads to infinity (Wang, 2017; Morgan, Banks and Carson, 1984).

3 MATHEMATICAL MODEL DEVELOPMENT

The development of the queuing model is a vital and challenging step in expressing the model into a mathematical equation because it needs to include all necessary variables. Let us consider that the model is based on FIFO queue discipline, inter-arrival and service are exponential, infinite population, and every customer waits to be served regardless of the queue length and $\lambda<c\mu$. Different studies used a mathematical equation to calculate the performance of the queue. Many researchers used an analytical model to analyze the queuing problem, similarly, in this paper, we used the formula of (Ragapriya, 2019; Ng Chee-Hock, 2008). Hence the traffic intensity with c servers expressed:

$$\rho = \lambda / c\mu \tag{1}$$

3.1 State balance equation of the Model

In this model, we examine the equilibrium (steady-state) in which P_n is independent of time as $T \rightarrow \infty$ $P_n(t)$ denoted by P_n, n=0,1,2…… hence the balance equation of M/M/c is expressed as (Ng Chee-Hock, 2008):

$$\lambda p_0 = \mu p_1 \tag{2}$$

$$(\lambda + c\mu) p_n = (n+1)\mu p_{n+1} + \lambda p_{n-1} \quad 1 \leq n < c \tag{3}$$

$$(\lambda + c\mu) p_n = c\mu p_{n+1} + \lambda p_{n-1} \qquad n \geq c \tag{4}$$

Based on the above equation, we have:

$$P_n = \begin{cases} p_0 \frac{(c\rho)^n}{n!}, \; n \le c \\ p_0 \frac{c^c \rho^n}{c!}, \; n \ge c \end{cases} \tag{5}$$

Therefore the normalization equation is expressed as follows:

$$\sum_{n=0}^{\infty} p_n = 1 \tag{6}$$

Now equation (1) using normalization equation gives

$$\mathrm{P}_0 = \left[\sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!} \frac{1}{1-\rho}\right]^{-1} \tag{7}$$

3.2 Performance analysis of M/M/c queuing system

After we know the p_0 and p_n we can analyze other queuing parameters of the system. The performance of the system is usually expressed in terms of a delay in the queue or in the system, which is expressed as the distribution of waiting time, distribution of the number of customers, distribution of idle and busy, etc. Different authors have measured the performance of the multi-server queue model almost in a similar way. This paper used the formulas presented in (Ragapriya,2019; Ng Chee-Hock, 2008 and Thomopoulos 2019):

The average number of vehicles in the toll service:

$$\mathrm{ES} = \sum_{n=0}^{C-1} n p_n + c \sum_{n=c}^{\infty} p_n = \frac{\lambda}{\mu} p_0{}^{-1} p_0 = \frac{\lambda}{\mu} \tag{8}$$

The average number of vehicles in the queue:

$$\mathrm{ELq} = \sum_{n=c+1}^{\infty} n p_n = p_0 \frac{(c\rho)^c}{c!(1-\rho)} \frac{\rho}{1-\rho} \tag{9}$$

The average number of vehicles in the system:

$$\mathrm{EL} = \sum_{n=1}^{\infty} n p_n = \left[p_0 \frac{(c\rho)^c}{c!(1-\rho)} \frac{\rho}{1-\rho}\right] + \frac{\lambda}{\mu} \text{ or ES+ELq} \tag{10}$$

Average waiting time of vehicles in the system:

$$\mathrm{ET} = \frac{EL}{\lambda} \tag{11}$$

The average waiting time of the vehicles in the queue:

$$\mathrm{EW} = \mathrm{ET} - 1/\mu = \mathrm{ELq}/\lambda \tag{12}$$

4 SIMULATION MODEL

Simulation modeling is a crucial phase in optimizing an analysis study. The simulation model stage is concerned with the development of an accurate simulation model that can imitate the behavior of the underlying system and recreate important findings for further inquiry. Simulations can be continuous simulation models that simulate systems with state variables that change continuously over time or discrete event simulation models that simulate systems with state variables that change only at a discrete point over time(Kin and Chan, 2011). Simulation modeling analyzes the behavior of the system using input variables running for a specific period. Research Carson and Maria (1997) indicated that in this phase of simulation development, the conceptual model is constructed from the real-world system and then programmed into the computer-based simulation model. This helps to iterate the model without the physical cost to get the optimal server of the toll service. Authors Ouzayd, Saadi and Benhra (2012) show the importance of the Colored Petri net simulation model in the medicine circuit to optimize service by balancing expenses and demands.

Different researchers use different computerized software to simulate the conceptualized model; however, in this paper, we use Colored Petri nets which allows us to test the validity of the mathematical model analysis. We applied the CPN tool software version 4.0.1 sourced from opened access (https://cpntools.org/2018/01/15/windows/). Developing the system using Colored Petri nets by the modeler helps to design, check, and analyze the model easily while for customers it is easy to understand as it is presented graphically (Hafilah et al., 2019).

Colored Petri nets (CPNs) are a graphical language for modeling and analyzing features of concurrent and distributed systems. CPNs add data types, functions, and modules to PNs(Guo et al., 2017). CPN model builds using place, transitions, and arcs combined with expressions.

Fig. 2 Conceptual model of the toll service; source: author

Before modeling the simulation model, first, we need the declaration of color sets, which is vital to the function of CPN models. Colors in CPN specify the type of the place or the type of token that a place can receive eg INT, Unit, String, Boolean, enumerated or index. As a result, colors are interpreted based on the color set declared on cpn model. As Fig. 3 presents, some declarations such as color set string, bool, real etc. are found in the CPN tool while others such as INTINFlist, UNITtm, var vehicles etc. are declared by the modeler.

```
▼Declarations
  ▶Standard priorities
  ▼Standard declarations
    ▼colset UNITtm = unit timed;
    ▼colset INTINF = intinf;
    ▼colset INTINFtm = intinf timed;
    ▼colset INTINFlist =list INTINF;
    ▼var vehicles: INTINFlist;
    ▼var vehicle: INTINF;
    ▼fun ET (EX) = round (exponential (1.0/EX));
```

Fig. 3 Declaration of the colourset; source: author

The CPN model shown in Fig. 4 has 6 places and four transitions. The place 'next' indicates the state of the system from the beginning which is connected to the transition 'vehicle arrival'. The arc expression from the transition 'vehicle arrival' to the next indicates that the vehicle arrived with the function exponential time (4.28 seconds). If the toll counters are busy, the arrived vehicle adds to an auxiliary place and waits until the server becomes free 'vehicle waiting for service' because it cannot add to the place queue as it has one token. If the counters are idle, the vehicle arrives immediately, does not queue, and gets the service. The vehicle arriving at the 'toll service' place 'toll service' gets service with the function exponential time (15 seconds). When the service finishes, the vehicle departs from the system and the free counter goes to the place 'idle server'. In addition to this arc expressions such as 'vehicles^^ [time ()]' and vehicle:: vehicles are used. The function time () returns the actual simulated time and operator '^^' concatenation of the list and '::' is used to express if the toll service is free and if there is at least one vehicle in the queue of vehicles.

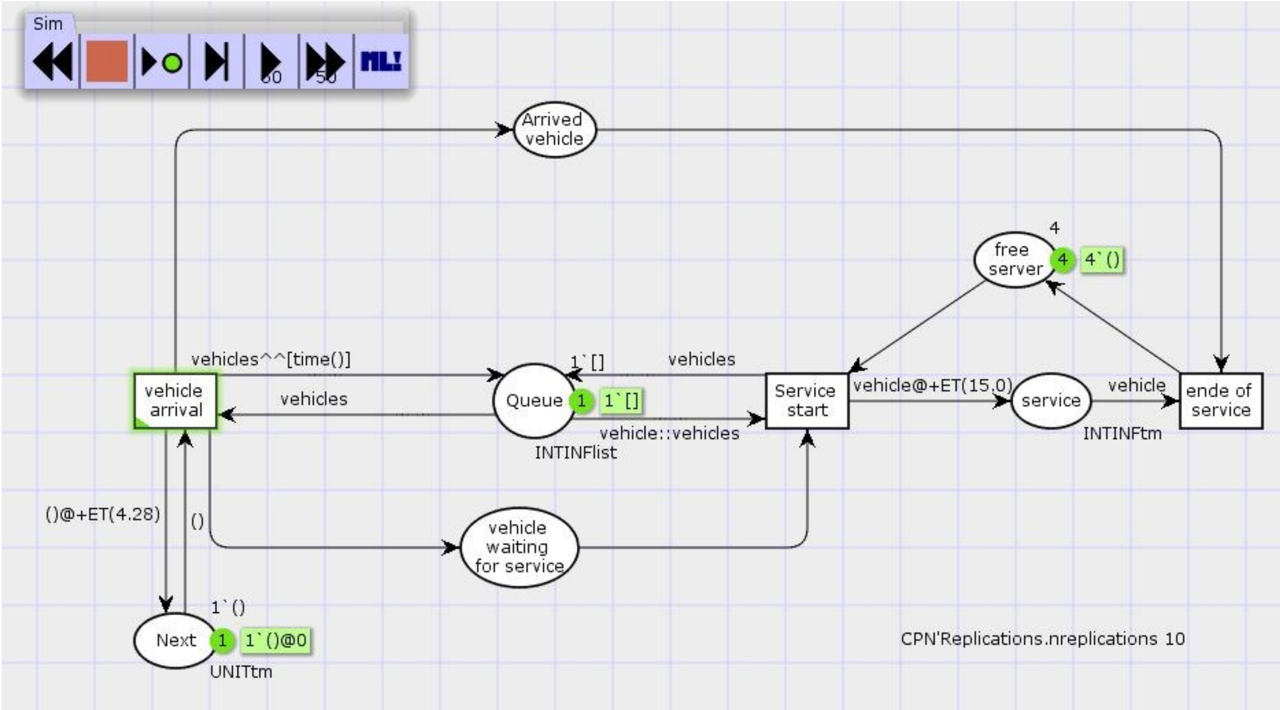

Fig. 4 Toll queue model using the CPN tool; source: author

The place expressed as a free server and service are used to state the counter as idle and busy, respectively. The free server indicates the server is ready to provide service whereas the place toll service indicates all servers are busy and it orders the vehicle must add to the place 'vehicle waiting for service'.

4.1 Performance analysis using a monitoring function

Based on this system and arc expression, we can measure the performance of the toll service using the monitor function. To analyze the performance of the system, we create 'ES' to monitor the mean number of vehicles in the service that is connected to the place 'service'. The monitor of 'EL' helps analyze the mean number of vehicles in the system by linking the place 'arrived vehicle'. The monitoring function 'ELq' analyzes the mean number of vehicles in the queue linked to the place 'vehicle waiting for service'.

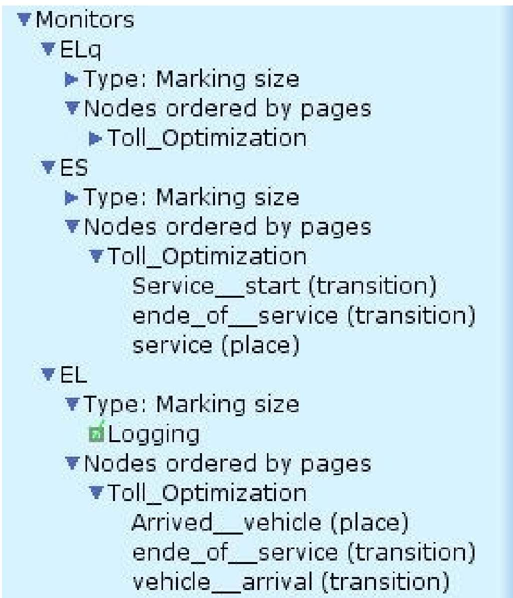

Fig. 5 monitoring function by mark size monitor; source: author

To monitor the meantime a vehicle waits in the queue, we use the monitoring function 'EWq' connected with the place 'queue'. Finally, we analyze the meantime the vehicle spent in the system using the monitoring function 'ET' linking with the transition 'end of service'. In this simulation, to terminate at the specific time, we use the breakpoint monitoring function 'fun pred = IntInf.toInt(time())>2592000'. Fig. 6 indicates the mean waiting time in the queue (EWq) and the mean waiting time in the system can be calculated by deducting vehicle from time.

```
▼EW
  ▶Type: Data collection
  ▶Nodes ordered by pages
  ▶Predicate
  ▼Observer
    fun obs (bindelem) =
    let
      fun obsBindElem (Toll_Optimization'Service_start (1,
                       {vehicle,vehicles})) = time()-vehicle
        | obsBindElem _ = ~1
    in
      obsBindElem bindelem
    end
  ▶Init function
  ▶Stop
▼ET
  ▶Type: Data collection
  ▶Nodes ordered by pages
  ▶Predicate
  ▼Observer
    fun obs (bindelem) =
    let
      fun obsBindElem (Toll_Optimization'ende_of_service (1, {vehicle})) = time()-vehicle
        | obsBindElem _ = ~1
    in
      obsBindElem bindelem
    end
  ▶Init function
  ▶Stop
```

Fig. 6 monitoring function of data collection monitor; source: author

5 RESULTS AND DISCUSSION

5.1 Experiment and result analysis

Using the above two models, we can analyze the performance of the Addis-Adama expressway toll service by considering on average 20000 vehicles served on this road. The mean-arrival rate of the vehicle is 828 per hour or 13.8 per minute and the service rate is 240 per hour or 4 per minute. Which means mean-arrival time of a vehicle is 4.28 seconds so the arrival rate is 0.23 vehicles per second. On this expressway, there are 10 toll servers with a mean service time of 15 seconds as the service is manual (road traffic technology, 2022). Hence, we used 10 counters starting from 4 to analyze the model to identify the optimal number of counters that balance waiting time and operating costs

As presented in Tables 5.1 and 5.2, the mathematical and simulation results indicate that the traffic intensity is 0.87, i.e. the utilization is 87% when the system uses four servers, which means the idleness of the server is 13%. On the other hand, when the system uses 10 servers, utilization drops to 35% and the idleness of the servers is 65%. This indicates that the mean number of vehicles in the service is low and the counter is out of a job when the number of counters is greater than seven servers. In the toll service, the numbers of vehicles in the queue declines as counters increase. The mean waiting time in the queue is 23.98 seconds (0.399 minutes) and then declines radically from four to the rest of the counter. The mean waiting time in the system was 38.88 seconds or 0.648 minutes on four servers after that it declined by 51% from the fourth counter to the fifth counters, however, the rest of the counters were almost constant with slight differences.

Tab. 1 Performance results from analytical analysis; source: author

Number of servers	Monitoring function				
	ES	ELq	EL	ET (sec)	EW (sec)
4	3.484848	5.516356	9.001204	38.88520128	23.98415
5	3.484848	0.915166	4.400014	19.00806048	3.978980
6	3.484848	0.255838	3.730686	16.11656352	1.112340
7	3.484848	0.072226	3.557074	15.36655968	0.314026
8	3.484848	0.025362	3.502100	15.1290720	0.110270
9	3.484848	0.003555	3.488403	15.03966096	0.015456
10	3.484848	0.001840	3.486688	15.01929216	0.007997

Tab. 1 and 2 indicates that the result obtained from both models is almost the same or lies within the confidence interval. Using 'CPN'Replications.nreplications 10' the result of the simulation is independent and identically distributed with a 95% confidence interval. Based on the output, it is bounded by lower and higher bounds. The basic issue in toll service is the mean waiting time of the vehicles in the system and queue. Based on this 'EW' becomes declines to a low level after eight servers but 'ET' after a decline from one to two and three servers in the rest of the servers becomes constant.

Tab. 2 Performance results from simulation analysis; source: author

Counters	Monitoring function									
	ES		ELq		EL		ET (sec)		EW (sec)	
4	8.76072	9.09179	5.25473	5.56997	3.48154	3.51181	37.50297	38.74991	23.49455	23.99961
5	4.39427	4.45823	0.88873	0.93814	3.48096	3.51181	18.78801	19.01178	3.799580	4.001590
6	3.72426	3.75157	0.25422	0.26829	3.48052	3.51584	16.08613	16.19382	1.087227	1.148010
7	3.54093	3.56419	0.07986	0.08341	3.48092	3.51077	15.32438	15.37910	0.301005	0.335782
8	3.49741	3.52348	0.02365	0.02600	3.48241	3.51866	15.06746	15.13372	0.101049	0.110831
9	3.47413	3.50536	0.00685	0.00834	3.48328	3.51781	14.99451	15.06560	0.014933	0.025647
10	3.47266	3.49959	0.00192	0.00236	3.48051	3.51751	14.97704	15.04157	0.007227	0.009009

5.2 Cost Analysis of Toll Service

The cost is generally divided into two direct costs which are related to tangible costs i.e. labor, equipment, commission, etc. and indirect costs are intangible related to waiting time costs. The difficult thing is calculating the cost incurred by waiting for the road user (Tiwari, Gupta and Joshi, 2016).

As previously discussed, the main goal of queuing theory is to determine the appropriate number of servers to deliver an efficient and effective service while minimizing service provider costs and customer waiting times. As a result, the operating performance of the toll queuing system is measured by a cost analysis of the service. In this analysis, we will use the data from the toll model to create a cost model for the queuing system under consideration. The customer's waiting cost in the system, as well as the service provider's operating costs, must be balanced. The customer's waiting cost is the amount of time it takes for them to receive service, from the time they arrive until the time they leave.

First, we calculate the total cost to establish the balance cost for the service system because the two costs become equal at a certain number of servers. We use the wage of a cashier as the basis for this paper, with a manual mechanism to collect fees taken as operating costs and vehicle waiting costs. Hence, the total cost is calculated as described by (Mitewu, 2019):

$$TC = EC^*EL + Cs^*C, \tag{13}$$

where TC - the total cost of the system, EC - Average waiting cost of each vehicle, EL - the average number of vehicles in the system, Cs - wages of the cashiers and C - number of cashiers.

Hence, to analyze the cost optimization, we use the estimated data rather than the real one i.e. that is not the real cost, rather it is based on a guesstimate because the cost fluctuates every time. So, assume that the company pays birr 500 per hour (0.0.139 per second) for each vehicle for waiting time and birr 300 per hour (0.0.0833 per second) for a wage of counters. Based on these data, we can analyze the cost of optimization. Note that other operating and waiting costs remain constant.

As tab. 3 indicated that the total cost decreased until using seven counters. However, when eight and more counters were used, the cost began to rise again. This showed that the optimal configuration is 7 counters. From the beginning, the total cost decreased from one to two counters by 73.56%. This means that increasing the number of counters can decrease the mean waiting time and number in the queue, however, the optimization started to decline which influenced the organization to lose. Similarly, the study by Tiwari, Gupta and Joshi (2016) showed that the number of servers changes with the cost of the system but at a certain point, the minimum expected cost is found.

Tab. 3 Result obtained by cost analysis of toll service; source: author

Number of counters (C)	**Cost analysis**				
	EL	**wage of a counter(Birr) per hour**	**EW(sec)**	**EC**	**TC**
4	9.001204	0.33320	23.98415	3.333797	3.66699685
5	4.400014	0.41650	3.978980	0.553078	0.96957822
6	3.730686	0.49980	1.112340	0.154615	0.65441526
7	3.557074	0.58319	0.314026	0.043650	0.62674961
8	3.502100	0.66640	0.110270	0.015328	0.68172753
9	3.488403	0.74970	0.015456	0.002148	0.75184838
10	3.486688	0.83300	0.007997	0.001112	0.83411158

Fig. 7 shows that the cost of the mean waiting time and the operating cost aligned before the mean waiting cost were higher than the operating cost. While after seven servers the operating cost becomes higher than the waiting cost. This means that using fewer or more than seven servers creates either congestion that is inconvenient and dissatisfactory for a road user or loss for a service provider.

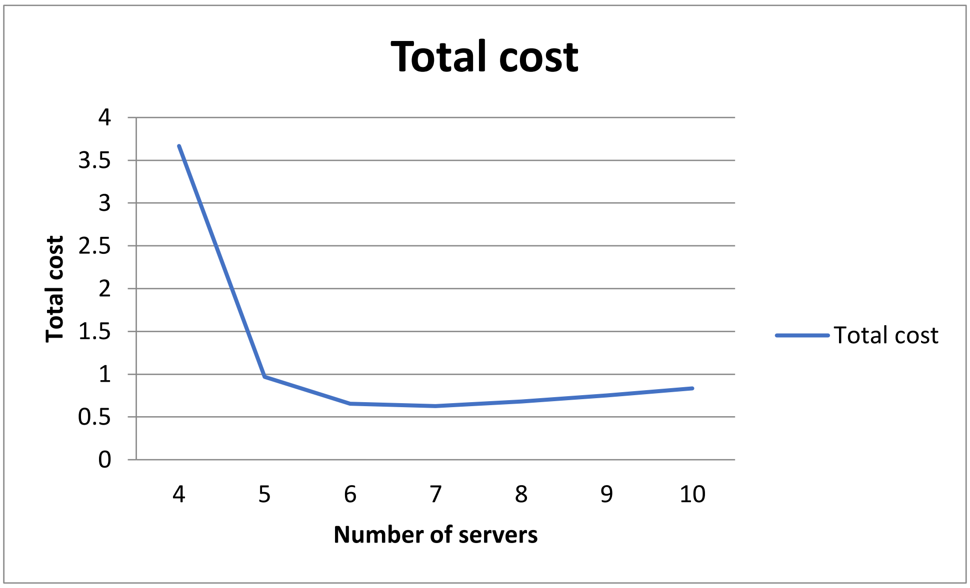

Fig. 7 The cost analysis result; source: author

6 CONCLUSIONS

This paper presents an improved optimization of the multi-server toll service model to balance the waiting time and service cost. As observed in the presented model and analysis, the toll service is determined by the number of vehicles in the queue, number of customers, service utilization, idleness of the system, and probability of the number of vehicles in the system. The study contributes to our understanding of the analysis of toll queues using mathematical models and the simulation using the CPN model. As the aim of the paper is optimization, the optimal number for a toll service must be configured with seven counters, which gives the maximum utilization of services based on the calibration of the above mean inter-arrival and service time.

Based on the result and analysis, before trying to solve the queue by increasing the number of servers, it must be analyzed based on the queue characteristics which indicate the optimal configuration of servers in a given system. Otherwise, it leads to loss because of the idleness of the servers. Generally, the randomness of inter-arrival and service time of the vehicle at toll services has made modeling toll service difficult. Therefore, this study has introduced a toll queuing model using CPN simulation and analytical analysis. The future work will be on the analysis of the vacation queuing model focusing on the probability of servers out of the main service line due to breakdown, repair, or other jobs.

Acknowledgements

The paper was supported by internal project of the Faculty of Mechanical Engineering, VSB – Technical University of Ostrava, SP 2022/62 Development and research in transport and logistics.

References

Abera, M. A. (**2019**). Evaluation of road traffic accident in Addis Ababa – Adama Expressway, *International Journal of Advance Research, Ideas and Innovations in Technology* 5(5), pp. 450–456.

Abubakar, U. M. *et al.* (**2016**). Application of Queuing Model in a Restaurant (A Case Study of Kalinga Restaurant), *International Journal of Science and Research (IJSR)*, 5(11), pp. 388–390. doi: 10.21275/nov163745.

Adeke, P. T. (**2018**). Modelling of queuing process at airport check-in system: a case study of Manchester and Leeds-Bradford airports, *Nigerian Journal of Technology*, 37(1), p. 35. doi: 10.4314/njt.v37i1.5.

Aksoy, G., Celikoglu, H. B. and Gedizlioglu, E. (**2014**). Analysis of Toll Queues by Micro-simulation: Results from a Case Study in Istanbul, *Procedia - Social and Behavioral Sciences*. Elsevier B.V., 111, pp. 614–623. doi: 10.1016/j.sbspro.2014.01.095.

Astutik, H. P. and Dewanti, D. (**2020**). The Effect of Toll Gate Type on the Queue of Vehicles in Connecting Roads: A case study of Bawen – Yogyakarta Toll Road, *Journal of the Civil Engineering Forum*, 6(1), p. 1. doi: 10.22146/jcef.43975.

Bastani, P. (**2009**). A queueing model of hospital congestion, *Simon Fraser University*, pp. 1–74.

Boahen, R. O., Adams, C. A. and Salifu, M. (**2013**). Performance evaluation of two toll Plazas on The Accra – Tema motorway, *International Refereed Journal of Engineering and Science (IRJES)*, 2(5), pp. 16–23.

BOARD, T. R. (**2006**). *NCHRP SYNTHESIS 364 Estimating Toll Road Demand and Revenue A*.

Carson, Y. and Maria, A. (**1997**). Simulation optimization: Methods and applications, *Winter Simulation Conference Proceedings*, pp. 118–126.

Cherng, C., Lewis, A. and Pai, S. (**2005**). Modeling toll plaza behavior using queuing theory, *Mathematical Contest in Modelling, University of Washington, Department of Mathematics, Washington, USA.*, p. 19. Available at: http://www.math.washington.edu/~morrow/mcm/cary05.pdf.

CPN tools. [Online]. Available at: https://cpntools.org/2018/01/15/windows/[accessed 2022, March 1].

Duhan, D. *et al.* (**2014**). Application of queuing theory to address traffic problems at a highway toll plaza, *Applied Mechanics and Materials*, 592–594 (December 2015), pp. 2583–2587. doi: 10.4028/www.scientific.net/AMM.592-594.2583.

Guo, Z. *et al.* (**2017**). A timed colored petri net simulation-based self-adaptive collaboration method for production-logistics systems, *Applied Sciences (Switzerland)*, 7(3). doi: 10.3390/app7030235.

Hafilah, D. L. *et al.* (**2019**). Modeling and simulation of air France baggage handling system with colored petri nets, *IFAC-papers online*. Elsevier Ltd, 52(13), pp. 2443–2448. doi: 10.1016/j.ifacol.2019.11.573.

Kin, W. and Chan, V. (**2011**). Foundations of Simulation Modeling, *Wiley Encyclopedia of Operations Research and Management Science*, (September). doi: 10.1002/9780470400531.eorms0336.

Krol, R. (**2018**). Tolling the Freeway: Congestion Pricing and the Economics of Managing Traffic, *SSRN Electronic Journal*. doi: 10.2139/ssrn.3211645.

Kumar, S. R., Vasudha, K., Rao, P. V. S. K., & Smitha, K. (**2019**). Performance Evaluation Of Toll Plazas Using Queuing Theory By, *i-manager's Journal on Civil Engineering*, 9(2), pp. 17–23.

Mckinnon, I. A., Knodler, M. A. and Christofa, E. (**2014**). Operational Analyzes of Varied Toll Plaza Configurations, *New England Section of the Institute of Transportation Engineers*, 54(3), pp. 11–16.

MITEWU, K. A. (**2019**). Modeling and Analysis of Queuing Systems in Banks: (A Case Study of Banque Commerciale du Congo), *Economic Policy*, Munich Per(16), pp. 0–33.

Molla, M. A.A. (**2017**). Case Study for Shuruchi Restaurant Queuing Model, *IOSR Journal of Business and Management*, 19(02), pp. 93–98. doi: 10.9790/487x-1902039398.

Morgan, C. B., Banks, J. and Carson, J. S. (**1984**). Discrete-Event System Simulation, *Technometrics*, 26(2), p. 195. doi: 10.2307/1268124.

Ng Chee-Hock, S.B.H. (**2008**). *Fundamentals With Applications in Communication Networks. 2nd Edition*.

Nico M. van Dijk Mark D. Hermans, Teunisse, Maurice J.G., H. S. and Ministry (no date). Designing The Westerscheldetunnel Toll Plaza Using A Combination Of Queueing And Simulation, *Proceedings of the 1999 Winter Simulation Conference*, pp. 1272–1279.

Ouzayd, F., Saadi, J. and Benhra, J. (**2012**). Proposed simulation models in Medicine drugs circuit with UML and colored petri net: Case Moroccan hospital system, *International Review on Modelling and Simulations*, 5(1), pp. 497–505. doi: 10.15866/ireit.v1i1.6370.

Prasad, S.V. *et al.* (**2018**). Mathematics And its Applications A Study on Multi Server Queuing Model to Optimize the Performance of a Toll Plaza Queuing System and Mathematical Model Analysis The basic indexes of the queuing systems, 6, pp. 355–359.

Ragapriya, M. (**2019**). Application of Queueing Theory in Engineering, VI(Iv), pp. 173–176.

Tiwari, S. K., Gupta, V. K. and Joshi, T. N. (**2016**). M/M/S Queueing Theory Model to Solve Waiting Line and to Minimize Estimated Total Cost, *International Journal of Science and Research (IJSR)*, 5(5), pp. 1901–1904. doi: 10.21275/v5i5.20051603.

Varghese, N. and Chandran, A. (**2021**). Application of Queuing Theory in Transportation, 9(6), pp. 55–58. Available at: www.ijert.org.

Wang, C. (**2017**). Study on Toll Plaza Design Based on M/M/1 Queue Theory, 61(Emcs), pp. 755–759. doi: 10.2991/emcs-17.2017.144.

Xiao, H. and Zhang, G. (**2010**). The queuing theory application in bank service optimization, IEEE, 2(072300410110), pp. 1097–1100. doi: 10.1109/ICLSIM.2010.5461127.

Yan, H. and Lam, W. H. K. (**1996**). Optimal road tolls under conditions of queueing and congestion, *Transportation Research Part A: Policy and Practice*, 30(5), pp. 319–332. doi: 10.1016/0965-8564(96)00003-1.

Road Traffic Technology© Copyright **2022**, Verdict Media Limited. Addis Adaba Adama (Nazareth)Expressway. (online).available at: https://www.roadtraffictechnology.com/projects/addis-adaba-adama-nazareth-expressway/ [accessed : 2022 ,September 22].