

Research Article

Steady-state probabilities of two queues in parallel

Leen Aburub¹, Ameen Alawneh², Aseel Aburub^{3*}

¹Department of Mathematics, University of Szeged, Szeged, 6722 Hungary.

²Department of Mathematics, Jadara University, Irbid, Jordan.

³Department of Physiotherapy, Applied Science Private University, Amman, Jordan.

ARTICLE INFO

Article history:

Received 3 Oct 2024

Accepted 11 Dec 2024

Published 30 Dec 2024

DOI: [10.35192/jjoas-n.v18i1.2020](https://doi.org/10.35192/jjoas-n.v18i1.2020)

*Corresponding author:

Department of Physiotherapy, Applied Science Private University, Amman, Jordan.

Email: a_abualrob@asu.edu.jo

Keywords:

Parallel queues.
Performance measures.
Steady state probability.
Queueing theory.

ABSTRACT

Background: This study empirically examines a system of two parallel queues, each served by an exponentially distributed server with service rates μ_1 and μ_2 , respectively. The arrival process follows a Poisson distribution with rate λ for both queues. Customers are directed to the queue with the shortest length; in the case of equal queue lengths, customers randomly select a queue with equal probability. The main focus of this study is on analyzing the steady-state probabilities of the two parallel queueing systems. After determining the performance measures, an application was studied to compare a one-queue, one-server system with the two parallel queues system to identify which is more efficient.

Methods: Two scenarios were explored in this research: one in which customers can switch between queues (queue hopping) and another where no switching is allowed. The steady-state probabilities of the system were analyzed. Wolfram Mathematica software was utilized to perform the computations and to find the transition matrix of the system.

Results: The study presents the steady-state probability distributions for the system, analyzing two cases: with and without queue hopping. These probabilities reveal the impact of allowing or restricting customer movement between queues on overall system performance. The two systems were compared to determine which one is more efficient.

Conclusion/Implications: The findings highlight the effect of the number of queues and the service rate on performance measures. The study also discusses two different cases of the two parallel queues system: when jockeying between queues is allowed and when it is not permitted. A command in Wolfram Mathematica is provided to demonstrate how the truncated transition matrix of the system can be applied.

Introduction

The two systems studied in this paper are similar to the systems studied by Haight (1) but if the two queues have the same length, then the new arrival customer will join the first queue in Haight's system, while in the systems in this research, they can join any of the queues with the same probability $1/2$. Zhao and Grassman (2) studied a system with a traffic intensity $\rho = \lambda/(2\mu)$ using a different numerical algorithm. Tarabia (3) examined two parallel queues with jockeying and restricted capacities by using the matrix-analytical technique. Adan and Resing (4) discussed analytic solution methods for queueing models with multiple waiting lines.

Borst, et. al (5) considered a parallel system of queues where the service rate of each queue depends on the number of customers in all of the queues. Wang and Locker (6) assumed that the customers arrival follows a Poisson process of rate 2ρ and the service times are identically independent random variables following an exponential distribution of unit rate. They show that the numerical solution can be found by matrix-geometric method and the analytical solution using the generating function. Guillemain, et. al (7) examined the two parallel queues with infinite servers. Yuvarani and Vijayalakshmi (8) analyzed a multistage batch arrival queue with different server interruptions with a second server replaces the main server when the interruptions occur. A real-life application of queueing theory studied by Mubarak R, et. al (9) is cross-polarization of antenna in satellite communication, where the cross-polarization is a standard used to measure the quality of lines that built with VSAT. Another real-life application of queueing theory is the traffic flow. Researchers like Yusuf et. al^a studied a recent development of smart traffic lights, which is a system where the traffic lights can intelligently control the signal time according to the traffic flow conditions to decrease the

congestion. The goal of this research is to find the steady state probabilities of the two parallel queues system and compare it with a single-server single-queue system, after analyzing two cases of the parallel queues system to understand how the systems works and test their efficiency in a given example. This helps the companies or the organizer of the queues to choose the most suitable system they need.

In this article, a queueing system with two servers were studied, where, each server has its own queue. Naming them server A , queue A , server B , and queue B , respectively. The service time at server A follows an exponential distribution with service rate μ_1 , and at server B is exponentially distributed with rate μ_2 . Arrivals to each queue follow a Poisson process with rate λ . Customers like to join the queue with the shortest length. Then, analyzed two cases of this system one of them gives the permission to the customers to change their queue if the other queue becomes shorter than the queue they have joined, and the other case is when jockeying is prohibited. The balance equations of both systems will be calculated by Wolfram Mathematica software (10). Also, found the transition probabilities of the system, then its steady state probabilities. Then, compared between the given two queues in parallel system with a single-server single-queue system ($M/M/1$) to see which of these systems is more efficient than the other, and used two methods: computing the traffic intensity, and the performance measures of the system, where the system with the least traffic intensity and the least performance measures, is the most efficient.

Methods

This study explored the simplest queueing system, $M/M/1$.

It started by understanding how the queuing system is built and which probability distributions to follow, where the arrivals rate follows a Poisson distribution with a parameter λ and the service rate follow an exponential distribution with parameter μ . Hence, the traffic intensity will be $\rho = \lambda/\mu$, and from this, the states probabilities and its performance measures of the system are found. Similarly, two parallel queuing system were analyzed and applied using two cases: when jockeying is allowed and when the customer is forced to stay in their queue. Then, the steady state probabilities of the two parallel queues system were found, and Wolfram Mathematica software was used to find the truncated transition matrix of the system. Lastly, a comparison was made between the single queue system and the parallel queues system to test which one is more efficient.

Results

A queuing system is a mathematical model that contains servers and queues. Where, a queue can only be generated when its server is busy serving a customer, and hence, queuing theory helps studying the customer's waiting time, the queue's length, the service rate, and the arrival rate of the customers. A customer may be chosen to get served in multiple cases: first-come first-served; last-come last-served; randomly; or maybe adding some conditions and priorities to some customers. This research focused on the first-come first-served case for each studied system. The study presents the steady-state probability distributions for the two cases (with and without queue hopping). These probabilities reveal the impact of allowing or restricting customer movement between queues on the overall system performance.

Single-Queue Single-Server System

Assume a queuing model with one server and one queue. The arrivals to the system follow a Poisson distribution with parameter λ . The service time of the server follows an Exponential distribution with parameter μ . This system is denoted by $M/M/1$. Let $\rho = \lambda/\mu$ be the traffic intensity of the system. The queuing system is said to be in its steady state if $\rho < 1$.

Blake(11) showed that if $P_n(t + \Delta t)$ denotes the probability that the system is in state n at time $t + \Delta t$, then

$$P_0 = 1 - \rho,$$

$$P_n = \rho^n (1 - \rho),$$

And the performance measures of the system are:
The average number of customers in the system:

$$L = E(N) = \frac{\rho}{1 - \rho}$$

The average number of customers in the queue

$$E(N_q) = \frac{\rho^2}{1 - \rho}$$

The mean waiting time in the queue $E(T_q)$

$$E(T_q) = \frac{\rho}{\mu(1 - \rho)}$$

The mean waiting time in the system

$$E(T) = \frac{1}{\mu - \lambda}$$

Two Queues in Parallel

Two queues in parallel system contain two servers, server A and server B, respectively. Each server has its own queue, queue A and queue B, respectively. This system has the following assumptions:

Arrivals follow a Poisson process with parameter λ for each queue.

The service time follow an exponential distribution with parameters μ_1 and μ_2 , respectively.

The system is said to be in its steady state if $\rho_1 = \frac{\lambda}{\mu_1} < 1$ and $\rho_2 = \frac{\lambda}{\mu_2} < 1$

Denote $X(t)$ and $Y(t)$ to be the lengths of queue A and queue B, respectively, at a time t . Assume a customer will join the system, if

$X(t) < Y(t)$ then they will join queue A, if $X(t) > Y(t)$ then they will join queue B, and if $X(t) = Y(t)$ then then will join any of the queues.

Jockeying is Allowed

In this case, the last customer in a queue is allowed to travel to the other queue if it is shorter than their current queue. Obviously, this induced three cases here, either $X(t) = Y(t)$, $X(t) = Y(t) - 1$, or $X(t) = Y(t) + 1$. Then, the following equations hold:

1. $-2\lambda p_{00} + \mu_2 p_{01} + \mu_1 p_{10} = 0$
2. $-(\lambda + \mu_1) p_{10} + \mu_2 p_{11} + \lambda p_{00} = 0$
3. $-(\lambda + \mu_2) p_{01} + \mu_1 p_{11} + \lambda p_{00} = 0$
4. $-(2\lambda + \mu_1 + \mu_2) p_{ii} + \lambda(p_{i,i-1} + p_{i-1,i}) + (\mu_1 + \mu_2)(p_{i+1,i} + p_{i,i+1}) = 0$
5. $-(\lambda + \mu_1 + \mu_2) p_{i+1,i} + \lambda p_{ii} + \mu_2 p_{i+1,i+1} = 0$
6. $-(\lambda + \mu_1 + \mu_2) p_{i,i+1} + \lambda p_{ii} + \mu_1 p_{i+1,i+1} = 0$

Solving the two linear systems of equations [1]-[3] and [4]-[6], respectively, for $p_{0,0}$, $p_{1,0}$, $p_{1,1}$, $p_{i+1,i}$, $p_{i,i+1}$, and $p_{i+1,i+1}$. By Wolfram Mathematica:

$$\text{Simplify[LinearSolve[}\begin{pmatrix} \mu_2 & \mu_1 & 0 \\ 0 & -(\lambda + \mu_1) & \mu_2 \\ -(\lambda + \mu_2) & 0 & \mu_1 \end{pmatrix} \cdot \begin{pmatrix} 2\lambda p_{00} \\ -\lambda p_{00} \\ -\lambda p_{00} \end{pmatrix}]$$

$$\left\{ \left\{ \frac{\lambda(2\lambda - (-2 + p_{00})\mu_1 + p_{00}\mu_2)}{\mu_2(2\lambda + \mu_1 + \mu_2)} \right\}, \left\{ \frac{\lambda(p_{00}(\mu_1 - \mu_2) + 2(\lambda + \mu_2))}{\mu_1(2\lambda + \mu_1 + \mu_2)} \right\} \right\}$$

$$\left\{ \frac{\lambda(\lambda(2\lambda - (-2 + p_{00})\mu_2) + \mu_1(2(\lambda + \mu_2) - p_{00}(\lambda + 2\mu_2)))}{\mu_1\mu_2(2\lambda + \mu_1 + \mu_2)} \right\}$$

$$\text{Simplify[LinearSolve[}\begin{pmatrix} \mu_1 + \mu_2 & \mu_1 + \mu_2 & 0 \\ -(\lambda + \mu_1 + \mu_2) & 0 & \mu_2 \\ 0 & -(\lambda + \mu_1 + \mu_2) & \mu_1 \end{pmatrix} \cdot \begin{pmatrix} (2\lambda + \mu_1 + \mu_2)p_{ii} - \lambda(p_{i,i-1} + p_{i-1,i}) \\ -\lambda p_{i,i} \\ -\lambda p_{i,i} \end{pmatrix}]$$

$$\left\{ \left\{ \frac{-\lambda(p_0 + p_{-1+ii})\mu_2(\lambda + \mu_1 + \mu_2) + p_{ii}(\mu_1^2(\lambda + \mu_2) + \mu_1\mu_2(3\lambda + 2\mu_2) + \mu_2(2\lambda^2 + 2\lambda\mu_2 + \mu_2^2))}{(\mu_1 + \mu_2)^2(\lambda + \mu_1 + \mu_2)} \right\}, \right.$$

$$\left. \left\{ \frac{1}{(\mu_1 + \mu_2)^2(\lambda + \mu_1 + \mu_2)} (-\lambda p_0\mu_1(\lambda + \mu_1 + \mu_2) - \lambda p_{-1+ii}\mu_1(\lambda + \mu_1 + \mu_2) + p_{ii}(\mu_1^3 + \lambda\mu_2^2 + 2\mu_1^2(\lambda + \mu_2) + \mu_1(2\lambda^2 + 3\lambda\mu_2 + \mu_2^2))) \right\} \right\}$$

$$\left\{ \frac{-\lambda p_0(\lambda + \mu_1 + \mu_2) - \lambda p_{-1+ii}(\lambda + \mu_1 + \mu_2) + p_{ii}(2\lambda^2 + \mu_1^2 + \lambda\mu_2 + \mu_2^2 + \mu_1(\lambda + 2\mu_2))}{(\mu_1 + \mu_2)^2} \right\}$$

Jockeying is Not Allowed

The customers of this system will stay in their queue until they get served and leave the system, inducing these equations:

$$p_{0,1+i} = \frac{-\mu_1 p_{1,i} - \lambda p_{0,-1+i} + (\lambda + \mu_2) p_{0,i}}{\mu_2}$$

$$p_{1+i,0} = \frac{-\lambda p_{-1+i,0} - \mu_2 p_{i,1} + (\lambda + \mu_1) p_{i,0}}{\mu_1}$$

$$p_{i,j} = \frac{\lambda(p_{i-1,j} + p_{i,j-1}) + \mu_1 p_{i+1,j} + \mu_2 p_{i,j+1}}{-\lambda - \mu_1 - \mu_2}$$

Steady State as A Limiting Distribution of Two-Dimensional State Space Markov Chain

Let i_A be the number of customers in queue A (including the one being served), and i_B be the number of customers in queue B (including the one being served), where $i_A \geq 0$ and $i_B \geq 0$. Assume that $P_{0,0}[(i_A, i_B) \rightarrow (j_A, j_B)]$ defines the transition probability from state (i_A, i_B) to state (j_A, j_B) . As Alawneh (12) showed, the transition probabilities of this Markov chains can be written as, where $i_A > 0$ and $i_B > 0$:

1. $P_{\infty}[(0,0) \rightarrow (1,0)] = \frac{1}{2}$
2. $P_{\infty}[(0,0) \rightarrow (0,1)] = \frac{1}{2}$
3. $P_{\infty}[(i_A, 0) \rightarrow (i_A, 1)] = \frac{\lambda}{\lambda + \mu_1}$
4. $P_{\infty}[(i_A, 0) \rightarrow (i_A - 1, 0)] = \frac{\mu_1}{\lambda + \mu_1}$
5. $P_{\infty}[(0, i_B) \rightarrow (1, i_B)] = \frac{\lambda}{\lambda + \mu_2}$
6. $P_{\infty}[(0, i_B) \rightarrow (0, i_B - 1)] = \frac{\mu_2}{\lambda + \mu_2}$
7. $P_{\infty}[(i_A, i_A) \rightarrow (i_A, i_A + 1)] = \frac{\lambda}{2\lambda + \mu_1 + \mu_2}$
8. $P_{\infty}[(i_A, i_A) \rightarrow (i_A + 1, i_A)] = \frac{\lambda}{2\lambda + \mu_1 + \mu_2}$
9. $P_{\infty}[(i_A, i_A) \rightarrow (i_A - 1, i_A)] = \frac{\mu_1}{2\lambda + \mu_1 + \mu_2}$
10. $P_{\infty}[(i_A, i_A) \rightarrow (i_A, i_A - 1)] = \frac{\mu_2}{2\lambda + \mu_1 + \mu_2}$
- If $i_B > i_A$
11. $P_{\infty}[(i_A, i_B) \rightarrow (i_A + 1, i_B)] = \frac{\lambda}{\lambda + \mu_1 + \mu_2}$
12. $P_{\infty}[(i_A, i_B) \rightarrow (i_A, i_B - 1)] = \frac{\mu_2}{\lambda + \mu_1 + \mu_2}$
13. $P_{\infty}[(i_A, i_B) \rightarrow (i_A - 1, i_B)] = \frac{\mu_1}{\lambda + \mu_1 + \mu_2}$
- If $i_B < i_A$
14. $P_{\infty}[(i_A, i_B) \rightarrow (i_A, i_B + 1)] = \frac{\lambda}{\lambda + \mu_1 + \mu_2}$
15. $P_{\infty}[(i_A, i_B) \rightarrow (i_A, i_B - 1)] = \frac{\mu_2}{\lambda + \mu_1 + \mu_2}$
16. $P_{\infty}[(i_A, i_B) \rightarrow (i_A - 1, i_B)] = \frac{\mu_1}{\lambda + \mu_1 + \mu_2}$

Which can be divided the state space of this Markov chain into the following four types:

1. If the first coordinate is zero, then the transition probability is

$$P_{\infty}[(0, i_B) \rightarrow (j_A, j_B)] = \begin{cases} \frac{\lambda}{\lambda + \mu_2} & ; j_B - i_B = 0 \\ \frac{\mu_2}{\lambda + \mu_2} & ; j_B - i_B = 1 \\ 0 & ; \text{otherwise} \end{cases} \quad (7)$$

2. If the second coordinate is zero, then the transition probability is:

$$P_{\infty}[(i_A, 0) \rightarrow (j_A, j_B)] = \begin{cases} \frac{\lambda}{\lambda + \mu_1} & ; j_A - i_A = 0 \\ \frac{\mu_1}{\lambda + \mu_1} & ; i_A - j_A = 1 \\ 0 & ; \text{otherwise} \end{cases} \quad (8)$$

3. If both coordinates are zeros, then the transition probability is:

$$P_{\infty}[(0,0) \rightarrow (j_A, j_B)] = \begin{cases} \frac{1}{2} & ; j_B = 1, j_A = 0 \\ \frac{1}{2} & ; j_B = 0, j_A = 1 \\ 0 & ; \text{otherwise} \end{cases} \quad (9)$$

4. If both coordinates are positive,

If the first coordinate equals the second coordinate, then the transition probability is:

$$P_{\infty}[(i_A, i_A) \rightarrow (j_A, j_B)] = \begin{cases} \frac{\lambda}{2\lambda + \mu_1 + \mu_2} & ; j_B - i_B = 1 \text{ and } j_A = i_A \\ \frac{\lambda}{2\lambda + \mu_1 + \mu_2} & ; j_B - i_B = 1 \text{ and } j_A = i_A \\ \frac{\mu_1}{2\lambda + \mu_1 + \mu_2} & ; i_A - j_A = 1 \text{ and } j_B = i_B \\ \frac{\mu_2}{2\lambda + \mu_1 + \mu_2} & ; i_A - j_A = 1 \text{ and } j_B = i_B \\ 0 & ; \text{otherwise} \end{cases} \quad (10)$$

If the first coordinate greater than the second coordinate, then the transition probability is:

$$P_{\infty}[(i_A, i_B) \rightarrow (j_A, j_B)] = \begin{cases} \frac{\mu_2}{\lambda + \mu_1 + \mu_2} & ; i_B - j_B = 1 \text{ and } j_A = i_A \\ \frac{\lambda}{\lambda + \mu_1 + \mu_2} & ; j_B - i_B = 1 \text{ and } j_A = i_A \\ \frac{\mu_1}{\lambda + \mu_1 + \mu_2} & ; i_A - j_A = 1 \text{ and } j_B = i_B \\ 0 & ; \text{otherwise} \end{cases} \quad (11)$$

If the second coordinate greater than the first coordinate, then the transition probability is:

$$P_{\infty}[(i_A, i_B) \rightarrow (j_A, j_B)] = \begin{cases} \frac{\mu_2}{\lambda + \mu_1 + \mu_2} & ; i_B - j_B = 1 \text{ and } j_A = i_A \\ \frac{\lambda}{\lambda + \mu_1 + \mu_2} & ; j_B - i_B = 1 \text{ and } j_A = i_A \\ \frac{\mu_1}{\lambda + \mu_1 + \mu_2} & ; i_A - j_A = 1 \text{ and } j_B = i_B \\ 0 & ; \text{otherwise} \end{cases} \quad (12)$$

Truncating these transition probabilities to a certain level n , such that $P_n[(i_A, i_B) \rightarrow (j_A, j_B)]$, and it can be defined as follows:

1. The first coordinate is zero. From the transition probability function in equation (7), for each positive integer n and $i_B = 1, 2, \dots, n$

$$P[(0, i_B) \rightarrow (j_A, n+1)] = 0$$

$$P[(0, i_B) \rightarrow (n+1, i_B)] = 0$$

Then, the truncated transition probability function is:

$$P_n[(0, i_B) \rightarrow (j_A, j_B)] = \begin{cases} P_{\infty}[(0, i_B) \rightarrow (j_A, j_B)] & 0 \leq i_B, j_A, j_B \leq n \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{\lambda}{\lambda + \mu_2} & ; j_B - i_B = 0, 0 \leq i_B, j_A, j_B \leq n \\ \frac{\mu_2}{\lambda + \mu_2} & ; i_B - j_B = 1, 0 \leq i_B, j_A, j_B \leq n \\ 0 & ; \text{otherwise} \end{cases}$$

2. The second coordinate is zero. From the transition probability function in equation (8), for each n and $i_A = 1, 2, \dots, n$

$$P[(i_A, 0) \rightarrow (n+1, j_B)] = 0$$

$$P[(i_A, 0) \rightarrow (j_A, n+1)] = 0$$

Then, the truncated transition probability function is:

$$P_n[(i_A, 0) \rightarrow (j_A, j_B)] = \begin{cases} P_{\infty}[(i_A, 0) \rightarrow (j_A, j_B)] & 0 \leq i_A, j_A, j_B \leq n \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{\lambda}{\lambda + \mu_1} & ; j_A - i_A = 0, 0 \leq i_A, j_A, j_B \leq n \\ \frac{\mu_1}{\lambda + \mu_1} & ; i_A - j_A = 1, 0 \leq i_A, j_A, j_B \leq n \\ 0 & ; \text{otherwise} \end{cases}$$

3. Both coordinates are zeros. Clearly, from the transition probability function (9), if any of the coordinates is greater than one, then the transition probability

$$P_n[(0,0) \rightarrow (j_A, j_B)] = \begin{cases} P_\infty[(0,0) \rightarrow (j_A, j_B)] & 0 \leq i_A, j_A, j_B \leq n \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{1}{2} & ; j_A = 0, j_B = 1, 0 \leq i_A, j_A, j_B \leq n \\ \frac{1}{2} & ; j_A = 1, j_B = 1, 0 \leq i_A, j_A, j_B \leq n \\ 0 & ; \text{otherwise} \end{cases}$$

For the last three truncated transition probability functions, the remaining probabilities are zeros when considering the first n states in the first n subsets of the state space. Thus,

$$\lim_{n \rightarrow \infty} P_n[(i_A, i_B) \rightarrow (j_A, j_B)] = P_\infty[(i_A, i_B) \rightarrow (j_A, j_B)]$$

4. When the state is in the interior of the state space (both coordinates are positive integers):

The first coordinate equals the second coordinate. From the transition probability function in equation (10), for each n and $i_A = 1, 2, \dots, n$

$$P[(i_A, i_A) \rightarrow (i_A + 1, j_A)] = 0$$

$$P[(i_A, i_A) \rightarrow (i_A - 1, j_B)] = 0$$

$$P[(n, n) \rightarrow (n + 1, n)] = 0$$

$$P[(n, n) \rightarrow (n, n + 1)] = 0$$

Hence, the rest of the probabilities are not zeros and the truncated transition probability function is

$$P_n[(i_A, i_A) \rightarrow (j_A, j_B)] = \begin{cases} P_\infty[(i_A, 0) \rightarrow (j_A, j_B)] & 0 \leq i_A, j_A, j_B \leq n \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{\lambda}{2\lambda + \mu_1 + \mu_2} & ; j_B - i_B = 1, i_A, i_B, j_A, j_B \leq n \text{ and } j_A = i_A \\ \frac{\lambda}{2\lambda + \mu_1 + \mu_2} & ; j_A - i_A = 1, i_A, i_B, j_A, j_B \leq n \text{ and } j_B = i_B \\ \frac{\mu_1}{2\lambda + \mu_1 + \mu_2} & ; i_A - j_A = 1, i_A, i_B, j_A, j_B \leq n \text{ and } j_B = i_B \\ \frac{\mu_2}{2\lambda + \mu_1 + \mu_2} & ; i_B - j_B = 1, i_A, i_B, j_A, j_B \leq n \text{ and } j_A = i_A \\ 0 & ; \text{otherwise} \end{cases}$$

The first coordinate is greater than the second coordinate. From the transition probability function in equation (11) $i_A = 1, 2, \dots, n$ and for each $i_B = 1, 2, \dots, n - 1$

$$P[(i_A, i_B) \rightarrow (n + 1, j_B)] = 0$$

$$P[(i_A, i_B) \rightarrow (i_A, n + 1)] = 0$$

inducing,

$$P_n[(i_A, i_B) \rightarrow (j_A, j_B)] = \begin{cases} P_\infty[(i_A, 0) \rightarrow (j_A, j_B)] & 0 \leq i_A, j_A, j_B \leq n \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{\mu_2}{\lambda + \mu_1 + \mu_2} & ; i_B - j_B = 1, i_A, i_B, j_A, j_B \leq n \text{ and } j_A = i_A \\ \frac{\lambda}{\lambda + \mu_1 + \mu_2} & ; j_B - i_B = 1, i_A, i_B, j_A, j_B \leq n \text{ and } j_A = i_A \\ \frac{\mu_1}{\lambda + \mu_1 + \mu_2} & ; i_A - j_A = 1, i_A, i_B, j_A, j_B \leq n \text{ and } j_B = i_B \\ 0 & ; \text{otherwise} \end{cases}$$

The second coordinate is greater than the second coordinate. From the transition probability function in equation (12) and for each $i_A = 1, 2, \dots, n - 1$ and $i_B = 1, 2, \dots, n$

$$P[(i_A, i_B) \rightarrow (n + 1, j_B)] = 0$$

$$P[(i_A, i_B) \rightarrow (i_A, n + 1)] = 0$$

Thus, getting:

$$P_n[(i_A, i_B) \rightarrow (j_A, j_B)] = \begin{cases} P_\infty[(i_A, 0) \rightarrow (j_A, j_B)] & 0 \leq i_A, j_A, j_B \leq n \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{\mu_2}{\lambda + \mu_1 + \mu_2} & ; i_B - j_B = 1, i_A, i_B, j_A, j_B \leq n \text{ and } j_A = i_A \\ \frac{\lambda}{\lambda + \mu_1 + \mu_2} & ; j_B - i_B = 1, i_A, i_B, j_A, j_B \leq n \text{ and } j_B = i_B \\ \frac{\mu_1}{\lambda + \mu_1 + \mu_2} & ; i_A - j_A = 1, i_A, i_B, j_A, j_B \leq n \text{ and } j_B = i_B \\ 0 & ; \text{otherwise} \end{cases}$$

Resulting:

$$\lim_{n \rightarrow \infty} P_n[(i_A, i_B) \rightarrow (j_A, j_B)] = P_\infty[(i_A, i_B) \rightarrow (j_A, j_B)]$$

Now, the limiting distribution $\pi = \pi P$, where P is the truncated transition matrix. The size of π is $1 \times (n + 1)^2$, and the size of the truncated transition matrix P is $(n + 1)^2 \times (n + 1)^2$. Hence, resulting a linear system of $(n + 1)^2 + 1$ equations which should be solved for π . One of these equations is neglected.

Using Wolfram Mathematica software, define the truncated transition probability as:

$$P[i_A, i_B, j_A, j_B] :=$$

$$\text{Piecewise}\left[\left\{\frac{\lambda}{\lambda + \mu_2}, j_A == 1 \& i_A == 0 \& i_B > 0 \& j_B - i_B == 0\right\},\right.$$

$$\left\{\frac{\mu_2}{\lambda + \mu_2}, i_A == 0 \& j_A == 0 \& i_B > 0 \& i_B - j_B == 1\right\},$$

$$\left\{\frac{\lambda}{\lambda + \mu_1}, i_B == 0 \& j_B == 1 \& i_A > 0 \& j_A - i_A == 0\right\},$$

$$\left\{\frac{\lambda}{\lambda + \mu_1}, i_B == 0 \& i_A > 0 \& i_A - j_A == 1 \& j_B == 0\right\},$$

$$\left\{\frac{1}{2}, i_A == 0 \& i_B == 0 \& j_A == 0 \& j_B == 1\right\},$$

$$\left\{\frac{1}{2}, i_A == 0 \& i_B == 0 \& j_A == 1 \& j_B == 0\right\},$$

$$\{ \lambda / (2\lambda + \mu_1 + \mu_2), i_A == i_B \& j_B - i_B == 1 \& j_A == i_A \& 0 < i_A \& 0 < i_B \& 0 < j_A \& 0 < j_B \},$$

$$\{ \lambda / (2\lambda + \mu_1 + \mu_2), i_A == i_B \& j_A - i_A == 1 \& j_B == i_B \& 0 < i_A \& 0 < i_B \& 0 < j_A \& 0 < j_B \},$$

$$\left\{\frac{\mu_1}{2\lambda + \mu_1 + \mu_2}, i_A == i_B \& i_A - j_A == 1 \& j_B == i_B \& 0 < i_A \& 0 < i_B \& 0 < j_A \& 0 < j_B \right\},$$

$$\left\{\frac{\mu_2}{2\lambda + \mu_1 + \mu_2}, i_A == i_B \& i_B - j_B == 1 \& j_A == i_A \& 0 < i_A \& 0 < i_B \right\},$$

$$\left\{\frac{\mu_1}{\lambda + \mu_1 + \mu_2}, i_A > i_B \& i_B - j_B == 1 \& j_A == i_A \& 0 < i_A \& 0 < i_B \right\},$$

$$\left\{\frac{\lambda}{\lambda + \mu_1 + \mu_2}, i_A > i_B \& j_B - i_B == 1 \& j_A == i_A \& 0 < i_A \& 0 < i_B \& 0 < j_A \& 0 < j_B \right\},$$

$$\left\{\frac{\mu_2}{\lambda + \mu_1 + \mu_2}, i_A > i_B \& i_A - j_A == 1 \& j_B == i_B \& 0 < i_A \& 0 < i_B \right\},$$

$$\left\{\frac{\mu_2}{\lambda + \mu_1 + \mu_2}, i_B > i_A \& i_B - j_B == 1 \& i_A == j_A \& 0 < i_A \& 0 < i_B \right\},$$

$$\left\{\frac{\lambda}{\lambda + \mu_1 + \mu_2}, i_B > i_A \& j_A - i_A == \& i_B == j_B \& 0 < i \& 0 < i_B \& 0 < j_A \& 0 < j_B \right\},$$

$$\{ \{ \frac{\mu_1}{\lambda + \mu_1 + \mu_2}, i_B > i_A \& i_A - j_A == 1 \& i_B == j_B \& 0 < i_A \& 0 < i_B \}, 0 \}$$

To put the values in a list which represents a transition matrix P , use the following command:

```
Partition[Flatten[Table[P[i1,i2,j1,j2],{i1,0,2},{i2,0,2},{j1,0,2},{j2,0,2}]],9]
```

$$\begin{aligned} & \left\{0, \frac{1}{2}, 0, \frac{1}{2}, 0, 0, 0, 0, 0\right\}, \left\{\frac{\mu_2}{\lambda + \mu_2}, 0, 0, 0, \frac{\lambda}{\lambda + \mu_2}, 0, 0, 0, 0\right\}, \\ & \left\{0, \frac{\mu_2}{\lambda + \mu_2}, 0, 0, 0, \frac{\lambda}{\lambda + \mu_2}, 0, 0, 0\right\}, \left\{\frac{\mu_1}{\lambda + \mu_1}, 0, 0, 0, \frac{\lambda}{\lambda + \mu_1}, 0, 0, 0, 0\right\}, \\ & \left\{0, \frac{\mu_1}{2\lambda + \mu_1 + \mu_2}, 0, \frac{\mu_2}{2\lambda + \mu_1 + \mu_2}, 0, \frac{\lambda}{2\lambda + \mu_1 + \mu_2}, 0, \frac{\lambda}{2\lambda + \mu_1 + \mu_2}, 0\right\}, \\ & \left\{0, 0, \frac{\mu_1}{\lambda + \mu_1 + \mu_2}, 0, \frac{\mu_2}{\lambda + \mu_1 + \mu_2}, 0, 0, 0, \frac{\lambda}{\lambda + \mu_1 + \mu_2}\right\}, \\ & \left\{0, 0, 0, \frac{\mu_1}{\lambda + \mu_1}, 0, 0, 0, \frac{\lambda}{\lambda + \mu_1}, 0\right\}, \\ & \left\{0, 0, 0, 0, \frac{\mu_1}{\lambda + \mu_1 + \mu_2}, 0, \frac{\mu_2}{\lambda + \mu_1 + \mu_2}, 0, \frac{\lambda}{\lambda + \mu_1 + \mu_2}\right\}, \\ & \left\{0, 0, 0, 0, 0, \frac{\mu_1}{2\lambda + \mu_1 + \mu_2}, 0, \frac{\mu_2}{2\lambda + \mu_1 + \mu_2}, 0\right\} \end{aligned}$$

Where each list represents a row in the truncated transition matrix.

Comparing Between the Two Queuing Systems

In this section, a comparison is made between two queues in parallel system of traffic intensities $\rho_1 = \lambda/\mu_1$ and $\rho_2 = \lambda/\mu_2$ with an $M/M/1$ system with traffic intensity $\rho_3 = (2\lambda)/(\mu_1 + \mu_2)$

using two different methods. The first method is by comparing the values of the traffic intensities of the given two systems, where the system with the least traffic intensity is the most efficient.

Since, in the two parallel queues system, $\rho_1 = \lambda/\mu_1$ and $\rho_2 = \lambda/\mu_2$. Let the average traffic intensity of the system be

$$\hat{\rho} = \frac{\rho_1 + \rho_2}{2} = \frac{\lambda(\mu_1 + \mu_2)}{2\mu_1\mu_2}$$

Assume an $M/M/1$ system with traffic intensity

$$\rho_3 = \frac{\lambda_2}{\mu_3} = \frac{2\lambda}{\mu_1 + \mu_2}$$

Claim that in this case, $M/M/1$ is more efficient, since

$$\frac{2\lambda}{\mu_1 + \mu_2} < \frac{\lambda(\mu_1 + \mu_2)}{2\mu_1\mu_2}$$

$$4\mu_1\mu_2 < (\mu_1 + \mu_2)^2$$

$$4\mu_1\mu_2 < \mu_1^2 + 2\mu_1\mu_2 + \mu_2^2$$

$$-2\mu_2^2 < \mu_1^2 - 2\mu_1\mu_2 - \mu_2^2$$

$$-2\mu_2^2 < (\mu_1 - \mu_2)^2$$

is true since $\mu_1, \mu_2, \lambda > 0$.

Application: In a store, consider the following two queuing systems: the first queuing system consists of two cashiers where each must serve an average of 20 customers in their five-hour day. One of them can serve an average of 8 customers per hour. Where, the other cashier can serve an average of 10 customers per hour. The other queuing model is a vending machine that serves an average of 18 customers per hour. Which of these two queuing models is more efficient than the other?

In the first queuing model:

$\lambda = 20/5 = 4$ customers/hour, $\mu_1 = 8$ customers/hour, and $\mu_2 = 10$ customers/hour. Then, $\rho_1 = 4/8 = 0.5$, $\rho_2 = 4/10 = 0.4$, and $\hat{\rho} = (0.5 + 0.4)/2 = 0.45$ In the other system:

$\lambda_1 = (2\lambda) = 8$, customers/hour, and $\mu_3 = (\mu_1 + \mu_2) = 18$, customers/hour. Then, $\rho_3 = 8/18 = 0.4\bar{4}$.

Hence, the given $M/M/1$ is more efficient.

Another way to compare between two systems is by calculating their performance measures. Since in the two queues in parallel system there are two servers each has its own queue, then divide this system into two $M/M/1$ systems, that is, let system 1 contains queue A and server A ,

and system 2 contains queue B and server B . System 1 is an $M/M/1$ system with $\rho_1 = \lambda/\mu_1$, and system 2 is an $M/M/1$ system with $\rho_2 = \lambda/\mu_2$. Hence, finding the performance measures of the two queues in parallel as:

The average number of customers in the system

$$\hat{E}(N) = \frac{E_1(N) + E_2(N)}{2} = \frac{\frac{\rho_1}{1-\rho_1} + \frac{\rho_2}{1-\rho_2}}{2}$$

The average number of customers in the queue

$$\hat{E}(N_q) = \frac{E_1(N_q) + E_2(N_q)}{2} = \frac{\frac{\rho_1^2}{1-\rho_1} + \frac{\rho_2^2}{1-\rho_2}}{2}$$

The average waiting time in the queues

$$\hat{E}(T_q) = \frac{E_1(T_q) + E_2(T_q)}{2} = \frac{\frac{\rho_1}{\mu_1(1-\rho_1)} + \frac{\rho_2}{\mu_2(1-\rho_2)}}{2}$$

The average waiting time in the system

$$\hat{E}(T) = \frac{E_1(T) + E_2(T)}{2} = \frac{\frac{1}{\mu_1 - \lambda} + \frac{1}{\mu_2 - \lambda}}{2}$$

Now, compare this two queues in parallel with an $M/M/1$ with the assumption that its traffic intensity $\rho_3 = (2\lambda)/(\mu_1 + \mu_2)$.

From the table 1, it is shown that the given $M/M/1$ system is more efficient than the given two queues in parallel system.

Table 1. comparing between two systems

Arrivals and service time rates	Performance Measures	Single-queue single-server ($\rho_3 = \frac{2\lambda}{\mu_1 + \mu_2}$)	Two queues in parallel ($\rho_1 = \frac{\lambda}{\mu_1}, \rho_2 = \frac{\lambda}{\mu_2}$)
$\lambda = 13,$ $\mu_1 = 14,$ $\mu_2 = 15$	Average number of customers in the system	8.666666666666666	9.75
	Average number of customers in the queue	7.7701149425287355	8.852380952380953
	Average waiting time in the queue	0.2679349980182323	0.6189002267573697
	Average waiting time in the system	0.3333333333333333	0.75
$\lambda = 10,$ $\mu_1 = 20,$ $\mu_2 = 21$	Average number of customers in the system	0.9523809523809523	0.9545454545454546
	Average number of customers in the queue	0.4645760743321719	0.46645021645021645
	Average waiting time in the queue	0.01133112376419932	0.02280715316429602
	Average number of customers in the system	0.04761904761904762	0.09545454545454546
$\lambda = 1,$ $\mu_1 = 9,$ $\mu_2 = 8$	Average number of customers in the system	0.13333333333333333	0.13392857142857142
	Average number of customers in the queue	0.01568627450980392	0.015873015873015872
	Average waiting time in the queue	0.00092272202998847	0.0018876763668430334
	Average number of customers in the system	0.06666666666666667	0.13392857142857142

Discussion

It is important to study queuing theory because it helps with describing the features of the queues and helps with providing the information that the service providers need for optimizing their queuing models. Mathematicians can formulate and analyze many problems as queuing systems to help the companies to build the most efficient possible queuing model that works properly, costs as less as possible, and decreases their customers waiting time. Therefore, they want to build a queuing system that is in its steady state where its traffic intensity is less than 1. The approach is to study the performance measures with the assumption that the system reaches its steady state.

In future studies, more complicated queuing systems such as the parallel tandem queuing network can be analyzed and considering many different possible cases and conditions. For example, whether the jockeying is allowed or not and compare them with other known queuing systems. In addition to finding more applications and real-life situations using queuing theory to develop their system.

This research analyzed two different queuing systems, which are the $M/M/1$ system with traffic intensity $\rho = \lambda/\mu$ and the two queues in parallel system with traffic intensities $\rho_1 = \lambda/\mu_1$ and $\rho_2 = \lambda/\mu_2$, respectively. Also, studied two different types of two queues in parallel system, where in the first type jockeying is allowed, and it is prohibited in the other. Then, found the steady state probabilities of the two queues in parallel. Finally, compared between this system and $M/M/1$, where the traffic intensity of the $M/M/1$ system is $\rho_3 = (2\lambda)/(\mu_1 + \mu_2)$, by two different ways.

The first way is by computing the average between ρ_1 and ρ_2 in the two parallel queues system, $\rho = \lambda/(\mu_1 + \mu_2)$ and in $M/M/1$, where the most efficient system is the one with the least traffic intensity.

The other way is by finding the performance measures for both systems, and clearly, the system with the least values of its performance measures is the most efficient.

In both ways, the $M/M/1$ system with the traffic intensity $\rho_3 = (2\lambda)/(\mu_1 + \mu_2)$ is better than the given two queues in parallel system.

The findings highlighted the effect of queue hopping on the efficiency and balance of the system. Allowing customer movement between queues can optimize service time under certain conditions, providing insights into queue management strategies for systems with parallel servers.

Conclusion

The findings highlighted the probabilities for the two distinct cases: when queue hopping is allowed and when the customer is forced to stay in one queue. It also studied the effect of adding an additional server or increasing the service rate of an existing server to compare between these systems and test which one is more efficient in a given application. The research also found the steady state probabilities of two queues in parallel system and found its transition matrix using Wolfram Mathematica software.

Competing interests

The authors declare no competing interest.

Authors contributions

L Aburub, developed the research concept; and designed the data collection process; analysis of the data, and drafted the study reports. A. Alawneh and A Aburub provided academic guidance and continuous supervision, and reviewed the study.

Data information

No data to be share.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not for profit sectors.

References

1. Haight FA. Two queues in parallel. *Biometrika*. 1958;45(3-4):401-10.
2. Zhao Y, Grassmann WK. A numerically stable algorithm for two server queue models. *Queueing Systems*. 1991;8:59-79.
3. Tarabia AM. Analysis of two queues in parallel with jockeying and restricted capacities. *Applied Mathematical Modelling*. 2008;32(5):802-10.
4. Adan IJ-BF, Boxma OJ, Resing JAC. Queueing models with multiple waiting lines. *Queueing Systems*. 2001;37:65-98.
5. Borst S, Jonckheere M, Leskelä L. Stability of parallel queueing systems with coupled service rates. *Discrete Event Dynamic Systems*. 2008;18:447-72.
6. Wang PP, Locker VF. Steady-state distributions of parallel queues. *INFOR: Information Systems and Operational Research*. 2001;39(1):89-106.
7. Mahalakshmi R, Kalyanaraman R, editors. Two server queues with system breakdown. *The 11th National Conference on Mathematical Techniques and Applications*; 2019.
8. Chandrasekaran Y. M [x]/G/1 multistage queue with stand-by server during main server's interruptions. *Indonesian Journal of Electrical Engineering and Computer Science*. 2018;11(1).
9. Mubarak R, Budiyo S, Wulandari P, Rahayu F, Adriansyah A, Alaydrus M. A queue theory in the cross-polarization of antenna in satellite communication. *Indonesian Journal of Electrical Engineering and Computer Science*. 2021;22:884-92.
10. Torrence BF, Torrence EA. *The Student's Introduction to Mathematica and the Wolfram Language*: Cambridge University Press; 2019.
11. Pfeiffer PE, Schum DA. *Introduction to applied probability*: Elsevier; 2014.
12. Alawneh AJ. *Approximating the stationary distribution of two-dimensional Markov chains*: New Mexico State University; 1995.