Optimizing Resource Allocation in *M/M/* 1/*N* Queues with Feedback, Discouraged Arrivals, and Reneging for Enhanced Service Delivery

Amit Kumar[1], Savita[1,*] and Chandra Shekhar[2]

[1]*Department of Mathematics, Chandigarh University, Mohali, Punjab 140413, India*
[2]*Department of Mathematics, Birla Institute of Technology and Science Pilani, Pilani Campus, Pilani, Rajasthan, 333 031, India*
E-mail: amitk251@gmail.com; savy84@gmail.com; chandrashekhar@pilani.bits-pilani.ac.in
[*]*Corresponding Author*

Received 06 February 2024; Accepted 18 April 2024

Abstract

This article presents a novel computational approach for analyzing $M/M/1/N$ queues with feedback, discouraged arrivals, and reneging, under the first-come, first-served (FCFS) discipline. We calculate explicit transient state probabilities and represent results using symmetric tridiagonal matrix eigenvalues. Through numerical simulations, we validate our method, providing practical insights for optimizing resource allocation. Our study contributes to both theory and application, advancing queueing theory and aiding decision-makers in improving service quality and resource management.

Keywords: Queueing models, discouraged arrivals, reneging, resource allocation, feedback, service quality.

Journal of Reliability and Statistical Studies, Vol. 17, Issue 1 (2024), 1–16.
doi: 10.13052/jrss0974-8024.1711

1 Introduction

Queuing systems serve as fundamental models for studying and optimizing the flow of entities through service systems. In real-world scenarios, service facilities often encounter discouraged arrivals and reneging behavior from impatient customers, leading to complex dynamics that challenge efficient resource allocation and service quality. Moreover, incorporating feedback mechanisms further complicates the analysis of such systems. This study addresses these challenges by introducing a novel computational technique for analyzing $M/M/1/N$ queues with feedback, reneging, and discouraged arrivals, all operating under the first-come, first-served (FCFS) discipline. Our approach focuses on deriving explicit transient-state probabilities, enabling us to gain insights into the system's behavior. The findings have practical implications for enhancing service delivery and resource allocation, bridging the gap between queueing theory and real-world applications.

In today's technology-driven landscape, the seamless operation of automated finite-capacity queueing systems is indispensable across various sectors, including power, communication, security, manufacturing, and more. These systems are inevitably exposed to erratic customer behaviors, which can disrupt efficiency. To address this challenge, we employ a queueing theoretical approach to analyze and enhance the efficiency of such systems. The study of queueing models has garnered considerable attention from many theorists (*cf.* [1–3]). Therefore, it is valuable to provide a comprehensive summary of the notable contributions in this field. Ammar et al. [4] investigate a matrix technique to compute transient state probabilities for the finite waiting space single-server queue. In the past, many research papers concerning the queueing model of feedback (*cf.* [3,5–7]). The single-server Markovian queue with discouraged arrivals, commonly found in daily queueing situations, has been extensively studied due to its finite waiting space. This waiting line is handy for modeling a computing facility that exclusively handles batch-job processing. When the facility is heavily utilized, job submissions are discouraged, and arrivals can be represented by a Poisson process with a state-dependent arrival rate. Regardless of the number of jobs in the system, the time taken to process each job is exponentially distributed with a constant service rate. Kumar [8], Medhi & Choudhury [9], Reynolds [10] and Kumar [11] have also investigated the discouraged arrivals queue, and Sharma [12] investigated the solution of a queueing model using a triangular matrix approach. Recently, Kumar [13] discovered a solution for the single-server multiple-vacation queue with discouragement by employing confluent hypergeometric functions.

The study of reneging is pivotal for enhancing customer satisfaction, optimizing resource allocation, and improving overall system efficiency. Moreover, businesses and organizations benefit from considering reneging rates when making decisions related to staffing, service level agreements, and other operational aspects. In essence, reneging plays a central role in shaping the perceived quality of service, influencing customer decisions, and guiding strategic choices for effective queueing system management. Yang's [14] study optimized patient admission and queueing control policies in overcrowded emergency departments, minimizing costs and considering premature discharge decisions and patient reneging behavior. Atar's [15] study on large-time behavior in many-server queues with reneging revealed unique invariant states in fluid equations corresponding to Dirac measures, providing insights into stationary distributions and system behavior. Logothetis et al. [16] study explored the effectiveness of the reneging option in strategic queueing systems with server on-off periods, challenging the notion that balking alone is sufficient. Economou et al. [17] study explored customer strategic behavior in join-or-balk dilemma queueing systems with server vacations/failures, highlighting the impact of reneging on social welfare and throughput, particularly in unstable systems.

The structure of this paper unfolds as follows. Section 2 provides a comprehensive depiction of the queueing model, accompanied by introducing a concise equation for calculating time-dependent probabilities expressed in terms of the characteristic value of a symmetric tridiagonal matrix. Moving to Section 3, we delve into the performance measure for queue models. In Section 4, our focus shifts to a sensitivity analysis achieved through numerical experiments. Finally, Sections 5 and 6 encapsulate our findings and insights, offering conclusive remarks on the implications and contributions of our research.

2 Time Dependent Probabilities

In this study, we explore a queueing system that models customer behavior in the finite capacity service systems. We examine a queueing system where customers arrive following an exponential distribution. Customers may join the queue, balk due to impatience, or renege if the wait is too long. The system operates on a 'First Come First Serve' basis, with servers attending to customers immediately if available. We aim to analyze how customer behavior impacts system efficiency. The pertinent notations and assumptions are described as follows:

The inter-arrival times of customers follow the exponential distribution with a mean rate of λ. Upon arrival, if customers find the server idle, they receive immediate service; otherwise, they may wait in the system. Let $X(t); t > 0$ denote the number of customers present in the system at time t. If a customer finds the server busy serving predecessor customers, they may balk from the system, exhibiting impatience with a state-dependent probability of $\frac{n}{n+1}$. While waiting in the system, a customer may also renege, showing impatience with the waiting time. The duration before reneging follows an exponential distribution with a mean rate of α. The server serves waiting customers following a First Come First Serve (FCFS) service discipline. The service time for each customer follows an exponential distribution with a mean rate of μ. Served customers may rejoin the system to complete unsatisfactory service with probability χ, or they may leave the system with probability $\bar{\chi} = (1 - \chi)$. The current study is particularly significant due to the assumption of a finite capacity of size N.

Let $\pi_n(t) = \Pr(X(t) = n)$ represent the probability that there are n customers in the system at time t, where $n = 0, 1, 2, \ldots, N$. Leveraging the characteristic of memorylessness, we derive the subsequent collection of Chapman-Kolmogorov differential-difference finite equations for the time-dependent probabilities $\pi_n(t)$ in the following manner:

$$\frac{\mathrm{d}\pi_0(t)}{\mathrm{d}t} = -\lambda\pi_0(t) + (1-\chi)\mu\pi_1(t), \tag{1}$$

$$\begin{aligned}\frac{\mathrm{d}\pi_n(t)}{\mathrm{d}t} =& -\left(\frac{\lambda}{n+1} + (1-\chi)\mu + (n-1)\alpha\right)\pi_n(t) + \left(\frac{\lambda}{n}\right)\pi_{n-1}(t)\\ &+ ((1-\chi)\mu + n\alpha)\pi_{n+1}(t), \quad 1 \le n < N-1,\end{aligned} \tag{2}$$

$$\frac{\mathrm{d}\pi_N(t)}{\mathrm{d}t} = -\left[(1-\chi)\mu + (N-1)\alpha\right]\pi_N(t) + \left(\frac{\lambda}{N}\right)\pi_{N-1}(t) \tag{3}$$

The initial conditions are $\pi_0(0) = 1$ and $\pi_n(0) = 0; n = 1, 2, \ldots N$.

2.1 Evaluation of Probabilities

Laplace transform is an integral transform that simplifies the process of solving linear differential equations by transforming them into algebraic equations, which are often easier to solve.We define the Laplace transformation of state probabilities as follows:

$$\psi_n(\theta) = \int_0^\infty \mathrm{e}^{-\theta t}\pi_n(t)\mathrm{d}t$$

Through the defined Laplace transform, the system of governing differential-difference equations is transformed as a system of linear equations as follows:

$$\theta\psi_0(\theta) - 1 = -\lambda\psi_0(\theta) + (1-(1-\chi))\mu\psi_1(\theta) \tag{4}$$

$$\begin{aligned}\theta\psi_n(\theta) = -\left(\frac{\lambda}{n+1} + (1-(1-\chi))\mu + (n-1)\alpha\right)\psi_n(\theta) + \left(\frac{\lambda}{n}\right)\psi_{n-1}(\theta) \\ + ((1-(1-\chi))\mu + n\alpha)\psi_{n+1}(\theta), \quad 1 \le n < N-1,\end{aligned} \tag{5}$$

$$\theta\psi_N(\theta) = -\left[(1-(1-\chi))\mu + (N-1)\alpha\right]\psi_N(\theta) + \left(\frac{\lambda}{N}\right)\tilde{\pi}_{N-1}(\theta) \tag{6}$$

The above system of equations can be represented in matrix form as follows:

$$\boldsymbol{\omega}\boldsymbol{\Psi} = \boldsymbol{\delta} \tag{7}$$

Vector $\boldsymbol{\Psi} = [\psi_0(\theta), \psi_1(\theta), \psi_2(\theta), \ldots, \psi_N(\theta)]'$ and vector $\boldsymbol{\delta} = [P_0(0), P_1(0), \ldots, P_N(0)]'$ are column vectors of order $N+1$, and matrix $\boldsymbol{\omega}$ is a tridiagonal square matrix of order $N+1$, where

$$\boldsymbol{\omega} = \begin{bmatrix} \theta+\lambda & \Theta & 0 & 0 & \cdots & 0 & 0 \\ \lambda & \theta+\frac{\lambda}{2}+\Theta & -\Theta-\alpha & 0 & \cdots & 0 & 0 \\ 0 & -\frac{\lambda}{2} & \theta+\frac{\lambda}{3}+\Theta+\alpha & -\Theta-2\alpha & \cdots & 0 & 0 \\ 0 & 0 & -\frac{\lambda}{3} & \theta+\frac{\lambda}{4}+\Theta+2\alpha & \cdots & 0 & 0 \\ 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & 0 & \cdots & -\frac{\lambda}{N} & \begin{matrix}\theta+\Theta \\ +(N-1)\alpha\end{matrix} \end{bmatrix}.$$

and $\Theta = (1-\chi)\mu$.

The matrix $\boldsymbol{\omega}$ can be easily transformed into a tridiagonal form by a diagonal matrix.

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ 0 & 0 & d_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_{N+1} \end{bmatrix}$$

with

$$d_1 = 1,$$
$$d_r = \prod_{k=1}^{r-1} \sqrt{k \left(\frac{\Theta + (k-1)\alpha}{\lambda} \right)^k}$$

and we get

$$\theta \mathbf{I} + \mathbf{B} = \mathbf{D} \boldsymbol{\omega} \mathbf{D}^{-1},$$

where,

$$\mathbf{B} = \begin{bmatrix} \lambda & -\sqrt{\lambda\Theta} & 0 & 0 & \cdots & 0 & 0 \\ -\sqrt{\lambda\Theta} & \frac{\lambda}{2} + \Theta & -\sqrt{\frac{\lambda(\Theta+\alpha)}{2}} & 0 & \cdots & 0 & 0 \\ 0 & -\sqrt{\frac{\lambda(\Theta+\alpha)}{2}} & \frac{\lambda}{3} + \Theta + \alpha & -\sqrt{\frac{\lambda(\Theta+2\alpha)}{3}} & \cdots & 0 & 0 \\ 0 & 0 & -\sqrt{\frac{\lambda(\Theta+2\alpha)}{3}} & \frac{\lambda}{4} + \Theta + 2\alpha & \cdots & & \\ 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & 0 & \cdots & \frac{\lambda}{N} + \Theta + (N-2)\alpha & -\sqrt{\frac{\lambda(\Theta+(N-1)\alpha)}{N}} \\ 0 & 0 & 0 & 0 & \cdots & -\sqrt{\frac{\lambda(\Theta+(N-1)\alpha)}{N}} & \Theta + (N-1)\alpha \end{bmatrix}.$$

As the symmetric diagonal matrix elements are the same as that of $\boldsymbol{\omega}$ and non-diagonal elements in the r^{th} row being $\sqrt{\lambda((\Theta) + \alpha(r-2))/(r-1)}$ and $\sqrt{\lambda((\Theta) + \alpha(r-1))/r}$ respectively. Considering $\boldsymbol{\omega}_r(\theta)$ and $\mathbf{B}_r(\theta)$ as square submatrices of size r extracted from the bottom right and top left of the matrix $\boldsymbol{\omega}$ respectively, where $P_r(\theta)$ and $Q_r(\theta)$ denote their determinants, both $P_r(\theta)$ and $Q_r(\theta)$ adhere to the specified difference equations.

$$P_r(\theta) - \left(\theta + \frac{\lambda}{N-r+2} + \Theta + (N-r)\alpha \right) P_{r-1}(\theta) + \left(\frac{\lambda(\Theta + (N-r+1)\alpha)}{N-r+2} \right) P_{r-2}(\theta) = 0$$

$$Q_r(\theta) - \left(\theta + \frac{\lambda}{r} + \Theta + (r-2)\alpha\right) Q_{r-1}(\theta)$$
$$+ \left(\frac{\lambda(\Theta + (r-2)\alpha)}{r-1}\right) Q_{r-2}(\theta) = 0, \quad 2 \le r \le N,$$

with the initial conditions

$$P_0(\theta) = 1 = Q_0(\theta),$$
$$P_1(\theta) = \theta + \Theta + (N-1)\alpha, \quad Q_1(\theta) = \theta + \lambda.$$

It can easily be shown that

$$C_{\mathrm{st}}(\theta) = \sqrt{\frac{(\lambda(\Theta + (s-2)\alpha))^{t-s}}{s(s-1)\cdots t}} \frac{P_{N-t}(\theta) Q_s(\theta)}{|\theta I + \mathbf{B}|}, \quad s < t$$
$$= \frac{P_{N-s}(\theta) Q_s(\theta)}{|\theta I + \mathbf{B}|}, \quad s = t$$
$$= \sqrt{\frac{(\lambda(\Theta + (t-2)\alpha))^{s-t}}{t(t+1)\cdots s}} \frac{P_{N-s}(\theta) Q_t(\theta)}{|\theta I + \mathbf{B}|}, \quad s > t,$$

where

$$(C_{\mathrm{st}}) = (\theta I + \mathbf{B})^{-1}$$

Using [18], Eq^n (7) can be written as

$$\mathbf{\Psi} = \boldsymbol{\omega}^{-1}\delta$$
$$= \mathbf{D}^{-1}(\theta \mathbf{I} + \mathbf{B})^{-1}\mathbf{D}\delta,$$
$$\psi_n(\theta) = \sum_{j=0}^{N} d_n^{-1} d_j C_{nj}(\theta) P_j(0)$$
$$= d_n^{-1} d_i C_{ni}(\theta).$$

Also,

$$d_n^{-1} d_i = \sqrt{\frac{(\Theta + (n-1)\alpha)^{i-n} n(n+1)\cdots i}{\lambda^{i-n}}}, \quad n < i$$
$$= 1, \quad n = i$$
$$= \sqrt{\frac{\lambda^{n-i}}{i(i+1)\cdots n(\Theta + (i-1)\alpha)^{n-i}}}, \quad n > i,$$

so we obtain

$$\begin{aligned}
\psi_n(\theta) &= \sqrt{[(\Theta+(n-1)\alpha)((\Theta+(n-2)\alpha)]^{i-n}}\frac{P_{N-i}(\theta)Q_n(\theta)}{|\theta I+\mathbf{B}|}, \quad n<i \\
&= \frac{P_{N-n}(\theta)Q_n(\theta)}{|\theta I+\mathbf{B}|}, \quad n=i \\
&= \frac{\lambda^{n-i}}{i(i+1)\cdots n}\sqrt{\left[\frac{\Theta+(i-2)\alpha}{\Theta+(i-1)\alpha}\right]^{n-i}}\frac{P_{N-n}(\theta)Q_i(\theta)}{|\theta I+\mathbf{B}|}, \\
&\quad n>i, 0\le i, n\le N
\end{aligned}$$

The characteristic value of the matrix **B** are real distinct and non-negative [19](one characteristic value being zero) .

Let $\alpha_m(m=0,1,2,\ldots,N)$ be an characteristic value of **B** with $\alpha_0=0$, then we can write

$$|\theta\mathbf{I}+\mathbf{B}| = \theta\prod_{m=1}^{N}(\theta+\alpha_m)\,.$$

A partial fraction expansion is then performed to give

$$\psi_n(\theta) = \frac{R_n}{\theta}+\sum_{m=i}^{N}\frac{\boldsymbol{\omega}_{nm}}{\theta+\alpha_m} \tag{8}$$

where

$$\begin{aligned}
R_n &= \frac{\lambda^n}{n!\prod_{g=1}^{n}(\Theta+(g-1)\alpha)}\left\{1+\sum_{k=1}^{N}\frac{\lambda^k}{k!\prod_{g=1}^{n}(\Theta+(k-1)\alpha)}\right\}^{-1}, \\
&\quad n=0,1,2,\ldots,N, \\
\boldsymbol{\omega}_{nm} &= \sqrt{[(\Theta+(n-1)\alpha)(\Theta+(n-2)\alpha)]^{i-n}} \\
&\quad\times\frac{P_{N-i}(-\alpha_m)\,Q_n(-\alpha_m)}{(-\alpha_m)\prod_{k=1,k\neq m}^{N}(\alpha_k-\alpha_m)}, \quad 0\le n\le i, \\
&= \lambda^{n-i}\sqrt{\left[\frac{\Theta+(i-2)\alpha}{\Theta+(i-1)\alpha}\right]^{n-i}}\frac{P_{N-m}(-\alpha_m)\,Q_i(-\alpha_m)}{(-\alpha_m)\prod_{k=1,k\neq m}^{N}(\alpha_k-\alpha_m)}, \\
&\quad i<n\le N.
\end{aligned}$$

Now inverting Equation 8, we get

$$\pi_n(t) = R_n + \sum_{m=1}^{N} \boldsymbol{\omega}_{nm} \mathrm{e}^{-\alpha_m t}.$$

3 Performance Measure

Systematic observations are crucial for understanding and improving the performance of systems, including queueing systems. The expected value of the number of customers in a queueing system is a key metric that provides insights into the system's behavior. This metric is often referred to as the "average queue length" or "average number of customers in the system."

- If $X(t)$ represents the random variable for queue length and $E[X(t)]$ denotes its expected value, then

$$E(X(t)) = \sum_{i=1}^{N} i * \pi_i + \sum_{i=1}^{r} i \sum_{m=1}^{N} \omega_{i*m} e^{-\alpha_m t} + \sum_{i=r+1}^{N} i \sum_{m=1}^{N} \omega_{i*m} e^{-\alpha_m t}, \quad n \le r \le N \tag{9}$$

- The throughput of the system is defined as the average number of served customers at time t and is expected to be

$$\tau(t) = \sum_{i=1}^{N} \mu \pi_i \tag{10}$$

- The expected delay time is defined as the quotient of the expected number of customers in the system and the system's throughput at time t.

$$E_d(t) = \frac{E_n(t)}{\tau(t)} \tag{11}$$

4 Numerical Illustration

The numerical results from various experiments conducted using MAPLE software on a computing system with hardware configuration, including an Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz processor and 16.0 GB of

RAM, are summarized in Figures 1, 2, and 3. Analytical methods were used to derive explicit expressions for the transient system size probabilities, and hence related expectation. However, for practical insights, it is essential to visually represent these probabilities. Therefore, numerical presentations are provided to enhance the understanding of the transient system size probabilities in real-world scenarios.

Figure 1 illustrates the time-dependent system size probabilities for $\lambda = 2.1$, $\mu = 5$, $\alpha = 0.4$, $(1-\chi) = 0.8$, and a system capacity of $N = 20$. As time (t) increases, the transient-state probabilities tend to stabilize, indicating a convergence to the steady-state. These variations validate the effectiveness of our novel approach for determining the state probabilities.

In Figure 1, the time-dependent system size probabilities are plotted for $N = 20$, $\lambda = 3$, $\mu = 7$, $\alpha = 0.5$, and $(1 - \chi) = 0.07$. The evident trend of the state probabilities converging to a steady state is observed here, validating the proposed computational approach.

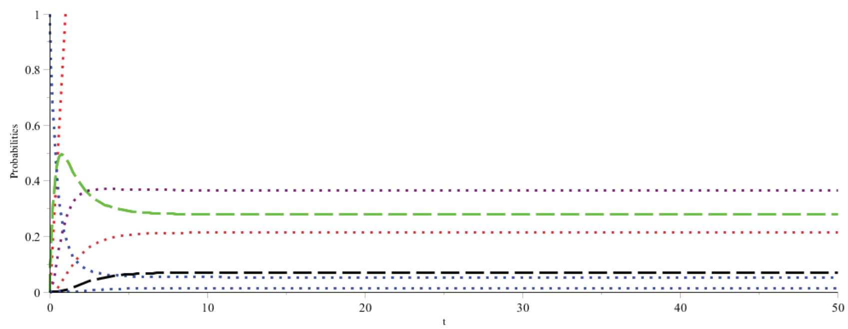

Figure 1 System size probabilities over time.

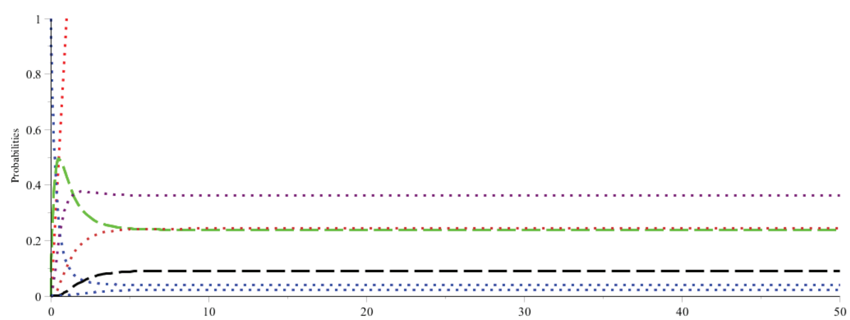

Figure 2 System size probabilities over time.

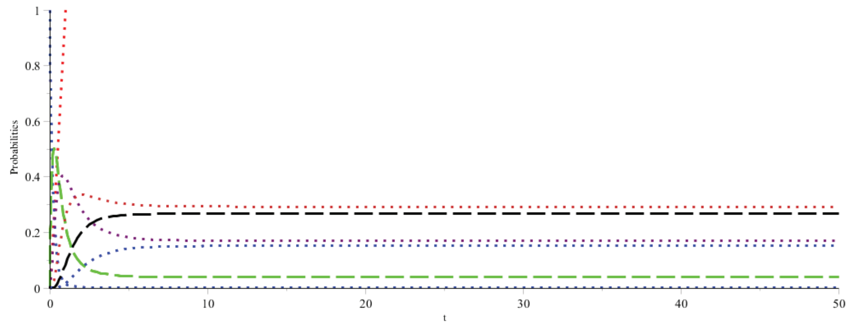

Figure 3 System size probabilities over time

In Figure 3 the time-dependent system size probabilities are plotted for $N = 20$, $\lambda = 7$, $\mu = 3$, $\alpha = 0.5$, and $(1 - \chi) = 0.07$. This graph also the results indicate that as the arrival rate λ increases, the system tends to be more congested, leading to higher probabilities of larger queue sizes. Conversely, increasing the service rate μ results in a faster processing of requests, reducing the queue size probabilities. The impact of discouraged arrivals, represented by the parameter α, is evident in the probabilities, especially as it approaches unity. The reneging parameter $(1 - \chi)$ also plays a significant role, affecting the overall queue dynamics.

5 Real Life Applications of This Model

In the realm of web server performance analysis, the single-server finite capacity queueing model proves to be exceptionally valuable for evaluating and enhancing the responsiveness of web servers. This model applies to scenarios where a web server receives a continuous stream of requests from users, each request varying in complexity. The server, however, can only process a limited number of requests concurrently, denoted as 'N'. In this intricate environment, several key factors influence the system:

- Users may arrive at the server at different rates, and their experience can be affected by congestion or delays, which can result in discouraged arrivals.
- Users may abandon their requests (reneging) if they perceive long waiting times.
- Feedback mechanisms may be implemented to dynamically adjust server resources based on traffic patterns and user demand.

The $M/M/1/N$ queueing model can be tailored to encompass these aspects for the analysis and enhancement of web server performance. It offers a systematic approach for modeling the arrival rate, service rate, and limitations on queue size. Discouraged arrivals can be accounted for by modifying the arrival rate to reflect factors influencing user satisfaction or server load. Reneging can be modeled as a probability, indicating the probability of users leaving the queue prematurely. Feedback mechanisms can be included to mimic dynamic adjustments in resource allocation in real-time.

6 Conclusion

This paper presents a novel and efficient computational technique for analyzing $M/M/1/N$ queues with feedback, discouraged arrivals, and reneging, particularly under the first-come, first-served (FCFS) discipline. We have derived explicit transient state probabilities, represented using symmetric tridiagonal matrix eigenvalues, emphasizing the model's significance. The validity of our approach is confirmed through numerical demonstrations, establishing its accuracy and providing practical insights for optimizing resource allocation in queueing systems. Our research is significant as it bridges the gap between theory and application, advancing queueing theory. By incorporating feedback mechanisms, addressing discouraged arrivals, and considering reneging scenarios, our model offers a more realistic representation of real-world queueing systems, enhancing its applicability to diverse service delivery environments.

7 Further Scope

The research conducted on $M/M/1/N$ queues with feedback, discouraged arrivals, and reneging lays a robust foundation for future exploration. To expand this work, fellow researchers could consider incorporating dynamic feedback mechanisms to enable adaptability to changing conditions. Additionally, extending the model to multi-class queueing systems could enhance its applicability by accommodating diverse customer characteristics. Furthermore, developing optimization algorithms based on the derived transient state probabilities could lead to real-time adjustments in resource allocation, particularly in response to varying demand patterns. These avenues offer promising directions for further research and application in queueing theory.

References

[1] C. Ancker Jr and A. V. Gafarian, "Some queuing problems with balking and reneging. I," *Operations Research*, vol. 11, no. 1, pp. 88–100, 1963.

[2] C. Ancker Jr and A. Gafarian, "Some queuing problems with balking and reneging—II," *Operations Research*, vol. 11, no. 6, pp. 928–937, 1963.

[3] C. Shekhar, A. Kumar, and S. Varshney, "Modified bessel series solution of the single server queueing model with feedback," *International Journal of Computing Science and Mathematics*, vol. 10, no. 3, pp. 313–326, 2019.

[4] S. Ammar, A. El-Sherbiny, S. El-Shehawy, and R. O. Al-Seedy, "A matrix approach for the transient solution of an m/m/1/n queue with discouraged arrivals and reneging," *International Journal of Computer Mathematics*, vol. 89, no. 4, pp. 482–491, 2012.

[5] R. Tolosana-Calasanz, J. Diaz-Montes, O. F. Rana, and M. Parashar, "Feedback-control & queueing theory-based resource management for streaming applications," *IEEE Transactions on parallel and distributed systems*, vol. 28, no. 4, pp. 1061–1075, 2016.

[6] E. G. Coffman and L. Kleinrock, "Feedback queueing models for time-shared systems," *Journal of the ACM (JACM)*, vol. 15, no. 4, pp. 549–576, 1968.

[7] C. Shekhar, A. Gupta, N. Kumar, A. Kumar, and S. Varshney, "Transient solution of multiple vacation queue with discouragement and feedback," *Scientia Iranica*, vol. 29, no. 5, pp. 2567–2577, 2022.

[8] R. Kumar, "A single-server markovian queuing system with discouraged arrivals and retention of reneged customers," *Yugoslav journal of operations research*, vol. 24, no. 1, 2016.

[9] P. Medhi and A. Choudhury, "Analysis of single server finite buffer queue under discouraged arrival and retention of reneging," *Yugoslav Journal of Operations Research*, vol. 33, no. 1, pp. 111–131, 2022.

[10] J. F. Reynolds, "The stationary solution of a multiserver queuing model with discouragement," *Operations research*, vol. 16, no. 1, pp. 64–71, 1968.

[11] R. Kumar and S. K. Sharma, "A multi-server markovian queueing system with discouraged arrivals and retention of reneged customers," *International Journal of Operations Research*, vol. 9, no. 4, pp. 173–184, 2012.

[12] O. Sharma, *Markovian Queues*. Allied Publishers, 1997.

[13] A. Kumar, "Single server multiple vacation queue with discouragement solve by confluent hypergeometric function," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 5, pp. 6411–6422, 2023.
[14] F. Yang, Q.-L. Li, C. Zhang, and C. Wang, "Optimal admission and queuing control with reneging behavior under premature discharge decisions," *International Transactions in Operational Research*, 2023.
[15] R. Atar, W. Kang, H. Kaspi, and K. Ramanan, "Long-time limit of nonlinearly coupled measure-valued equations that model many-server queues with reneging," *SIAM Journal on Mathematical Analysis*, vol. 55, no. 6, pp. 7189–7239, 2023.
[16] D. Logothetis, A. Manou, and A. Economou, "The impact of reneging on a fluid on-off queue with strategic customers," *Annals of Operations Research*, pp. 1–19, 2022.
[17] A. Economou, D. Logothetis, and A. Manou, "The value of reneging for strategic customers in queueing systems with server vacations/failures," *European Journal of Operational Research*, vol. 299, no. 3, pp. 960–976, 2022.
[18] J. W. Lewis, "Inversion of tridiagonal matrices," *Numerische Mathematik*, vol. 38, pp. 333–345, 1982.
[19] D. Yang and R. Gregory, "A survey of numerical mathematics, vol. ii," *Reading, Addison-Wesley*, p. 1097, 1973.

Biographies

Amit Kumar is an Assistant Professor at the University Institute of Science, Chandigarh University, India. He holds a Ph.D. from Birla Institute of Technology and Science, Pilani Campus, Rajasthan. His research interests encompass queueing theory, machine repair problem, optimal control, reliability and maintainability, stochastic modeling, sensitivity analysis, evolutionary computation, statistical analysis, and fuzzy set and logic. Kumar has published numerous research articles in reputable journals such as Reliability Engineering and System Safety, Journal of Computational and Applied Mathematics, Quality Technology and Quantitative Management, and Arabian Journal of Science and Engineering. He actively participates in conferences, Faculty Development Programs (FDPs), workshops, and symposiums as both a presenter and invited speaker. Kumar also serves as a reviewer for several prestigious journals and has professional experience visiting the irrigation department in Roorkee, India. (ORCID 0000-0001-5347-1808).

Savita is a research scholar in the Department of Mathematics at Chandigarh University. She earned her M.Sc. degree from GJU Hisar in 2006. Her primary research interests lie in the field of queueing theory and stochastic processes. She has actively participated in several national and international conferences.

Chandra Shekhar is a distinguished Professor and former Head of the Department of Mathematics at BITS Pilani, India. His research interests span queueing theory, computer and communication systems, machine repair problems, reliability and maintainability, stochastic processes, evolutionary computation, statistical analysis, and fuzzy set and logic. Professionally, he actively participates in academic events, presenting papers and delivering invited talks at national and international conferences and Faculty Development Programs (FDPs). He has organized multiple conferences, workshops, and symposiums and received recognition with the best research paper award at an international conference. He has authored over 50 research articles in prestigious journals, supervised three Ph.D. theses, and contributed to book chapters and textbooks. He serves as a member of editorial boards, a reviewer for reputable journals, and participates in various academic committees. He has professional experience working with esteemed organizations, including IIRS (ISRO), CSIR-IIP, NIH, WIHG, CPWD, NTPC, Bank of Maharashtra, and APS Lifetech. (ORCID: 0000-0002-2114-9096)