PERFORMANCE ANALYSIS OF THRESHOLD-BASED SCHEDULING SYSTEMS THROUGH DISCRETE TIME SYSTEM

ISSN (e) 2520-7393
ISSN (p) 2521-5027
www.estirj.com

Mahnoor Memon[1], Wajiha Shah[2], Qudsia Memon[3]

[1,2,3]*Department of Electronic Engineering, Mehran University of Engineering and Technology, Jamshoro*

Abstract: Most of the queueing systems are affected by time-dependent changes in system properties, including the response time or the service times. Inbound call volumes and agents' cyclic communication volumes in computer systems, time delayed air traffic at airports, customer queues in banks, and time-dependent truck arrival rates at seaports are highly congested systems. The purpose of the service provider is to find the lowest-cost admission and capacity allocation strategy so that they may dynamically decide when and who to serve. Threshold based scheduling system have been developed to model this situation. Analytical and simulation both methods are used for accuracy and validity purpose. The proposed system served as a best alternative solution to overcome the congestion problem in time scheduling systems.

Keywords: *Discrete-time queue; Threshold-based system, Markov chain, Matrix Geometric Method, Queuing system.*

1. Introduction

Customer satisfaction through improved quality of service is one of the main factors affecting a company's performance in today's highly competitive market. Transport networks, healthcare management systems, cloud - based services, system software, digital assistants, as well as other technologies are now becoming widespread in establishing successful scheduling mechanisms for serving incoming tasks employing in various ways. When it comes to the provision of high-quality service, large organizations must emphasize client satisfaction while optimizing efficiency and reducing costs. Service organizations have focused on a number of techniques to understand consumer perspectives and have produced a variety of tactics to provide better service to customers. The service business is particularly concerned with the length of time it takes for customers to receive service. Due to delays in receiving service, queuing is a typical problem in industrial settings and in ordinary life circumstances. A conventional queuing system has several critical elements, including line layout, demand groupings, arrival and service activities, and queue discipline [1]. Arriving entities are allocated the best available services, and if there aren't enough servers to serve everyone at once, waiting lines (buffers) are used. External and internal thresholds are used in generic scheduling systems. Threshold-based systems are useful when both system performance and cost are relevant. The system prioritizes serving jobs with the greatest threshold value in the queue[2], [3]. In this paper performance analysis of threshold-based scheduling systems through a discrete-time system is discussed, in which data is sampled and examined primarily using differential equations and the Markov chain flow method, which exhibits the model parameters.

2. Related Work

Many scholars have focused their attention on queueing systems in past decades. A single-class model based on probability-generating functions was used to evaluate system performance (system content and transmission delays) by applying preemptive and non-preemptive priority scheduling strategies [2]. Analysis of the system contents with various classes was accomplished by two Gauss Markov-modulated arrival processes [2], [4]. A classification technique categorizing approaches based on models with piecewise constant parameters, changed system properties, numerical and analytical solutions, and their underlying core principles were discussed in [5]. The simulation-based system indicating delays of queuing model for each customer associated with a random variable and the subsequent service times was implemented in [1], [6]. The article [7] addresses a discrete-time queueing system that provides a single server system with a deterministic (same event for every arrival) service rate. Three threshold levels namely (the scheduling threshold, maintenance threshold, and a failure threshold) were utilized to develop a condition-based maintenance (CBM) system [8]. The system's demonstration was based on the gamma process, which resulted in a better maintenance plan. In [9]analysis was carried out to model a video buffer based on G1/G1/1 queue with pq policy, ensuring the quality of downloading and playback videos simultaneously, which minimized the network bandwidth loss while streaming videos. The battery-operated sensors with energy harvesting capabilities

Corresponding author Email address*: mahnoormemonn@gmail.com*

were used to investigate transmission scheduling challenges for a remote state estimation problem in [10], by scheduling and transmitting local state estimates to a remote estimator across a packet dropping channel for minimizing estimation error covariance. Later, a discrete-time linear system was studied in [11]with partial information was studied by applying an event-triggered control strategy to achieve impressive performance gains. The Markov chain approach was used to assess a discrete-time queuing communication model and construct closed-form expressions for the system's process parameter [12]. The huge number of arrivals were handled in this study utilizing a single queue, single server approach based on the First come first serve discipline. Renewal theory and a stochastic process were employed to avoid system congestion throughout the repair procedure [13]. The paper [14] investigates how collected energy can be used in a wireless sensor network by putting a threshold on the system's token buffer to ensure the transmission of high-priority data packets. Since each of the methodologies mentioned has its own benefits but obtaining an accurate queuing-based system remains a challenging task. Hence, a queuing mechanism based on internal and external criteria has been implemented, to avoid system congestion. Threshold-based scheduling systems have received a lot of attention in the literature due to their accuracy and resilience [6], [8], [14]– [16].

3. Methodology

The flowchart of the overall system methodology is illustrated in the following Figure.1.

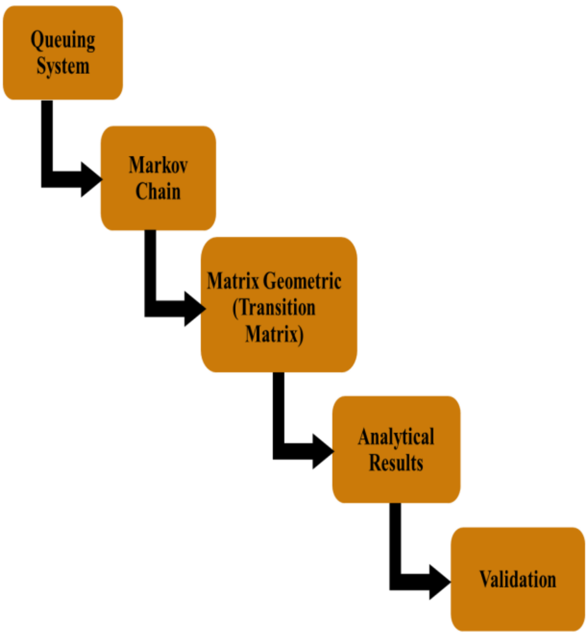

Figure.1. Flow chart of the developed methodology

3.1. Queueing System:

Queueing systems represent fundamental mathematical tools for explaining transportation congestion. When 'customers' demand 'service,' a queueing system emerges; in most cases, both the customers' arrival and the service hours are supposed to be random. Because the 'servers' are all filled, new customers will have to wait for the next available server. Figure.2 shows the basic structure of queueing process. An input source generates customers who require service over time. These consumers join a (finite or endless) queue after entering the queueing system. The queue discipline is a rule that selects a member of the queue for service at specific times (e.g., first-come-first-served). Afterward, the customer's desired service is provided by their service mechanism, and the customer departs the queueing system.

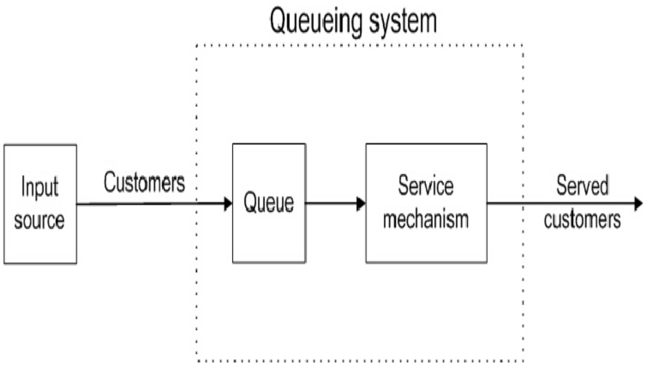

Figure.2. Block Diagram of Queueing management system

3.2. Markov Chain:

A mathematical system undergoing transitions between one state to another, depending on probability criteria. It has a distinguishing feature that all possible future states are fixed, regardless of how the process arrived at its current state.

3.2.1. Markov Chain for 2-D Queuing system:

The Markov chain for two-Dimensional queueing system is shown in Figure.3. The number of consumers in the queue corresponds to each state in the chain, and state transitions occur when new customers join the line or when customers finish their service and leave. The diagram represents that there is no arrival and no service in the mechanism at first. The Q1 increases horizontally while Q2 remain same, then Q2 increases vertically, Q1 remains same. Alpha H ($\propto_H$) and Alpha L ($\propto_L$) illustrate customer or data in queue one (Q1), Alpha L ($\propto_L$) depicts customer or data in queue two (Q2), while beta one (β_1) represents service and beta two (β_2) indicates that there is no service in the system.

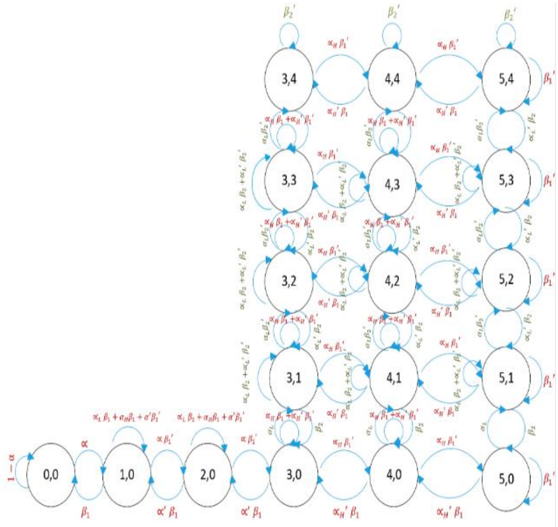

Figure.3. The 2-D Markov Chain

3.3. The Matrix Geometric Method (MGM):

The matrix geometric approach is a continuous-time Markov chain analysis method that uses transition rate matrices with a repeated block structure to analyze quasi-birth–death phenomena. The method implies a transition rate matrix with the tridiagonal block structure. The transition matrix using matrix geometric method for proposed system is shown in Figure.4.

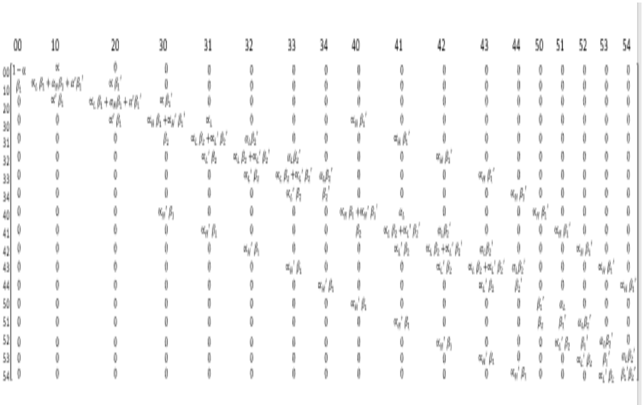

Figure.4. Finite Generator Matrix using Matrix Geometric Method

The implemented system is categorized into two sub-systems as shown in Figure.5.

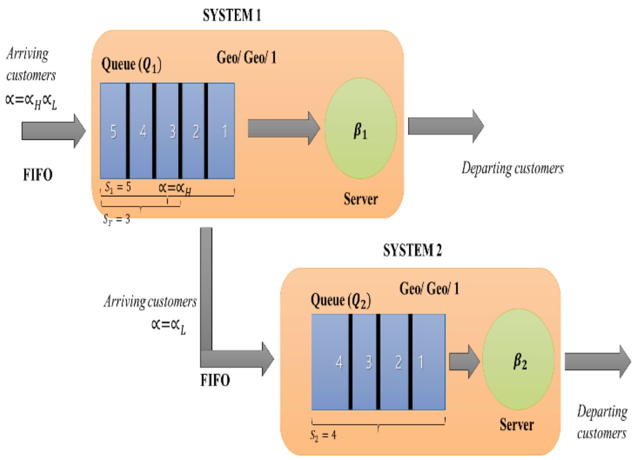

Figure.5. Block diagram of the implemented system

where $\propto_H$= high priority arrivals, $\propto_L$= low priority arrivals, Q_1 serves: $\propto_H$ and $\propto_L$, Q_2 serves: $\propto_L$, β_1= Q_1 server, β_2= Q_2 server and Limits for the threshold are $S_1 = 5, S_T = 3, S_2 = 4$. In the system1, the first queue Q_1, allows both high and low priority consumers till the threshold level S_T. Only high priority entries are accepted in Q_1, while low priority arrivals are routed to Q_2 in order to reduce congestion. When the S_T threshold is achieved, the first queue Q_1 only admits $\propto_H$ customers, while $\propto_L$ customers move to Q_2 to avoid the congestion. $\propto_H$ arrivals are serviced via Q_1 at the rate β_1 whereas $\propto_L$ arrivals are served by β_1 if no $\propto_H$ arrivals are available in Q_1 and by β_2.

The Markov chain and Finite generator matrix for system1 are shown in Figure.6 and 7 respectively.

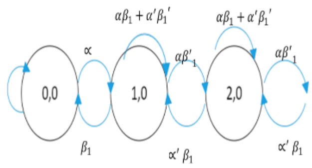

Figure.6. Markov chain for system 1

$$\begin{array}{c} \\ 00 \\ 10 \\ 20 \end{array}\begin{bmatrix} 00 & 10 & 20 \\ 1-\propto & \propto & 0 \\ \beta_1 & \alpha\beta_1+\alpha'\beta_1' & \alpha\beta'_1 \\ 0 & \propto'\beta_1 & \alpha\beta_1+\alpha'\beta_1' \end{bmatrix}$$

Figure.7. Finite Generator Matrix for system 1

Where $B_{00} = [1-\propto]$, $B_{01} = [\propto]$, and $B_{10} = [\beta_1]$ are Boundary and Border Matrices, $A_0 = [\propto \beta_1]$, $A_1 = [\propto \beta_1 + \propto' \beta_2]$, and $A_2 = [\propto' \beta_1]$ are Repetitive Matrices.

Similarly, the Finite generator matrix for system 2 is shown in Figure.8.

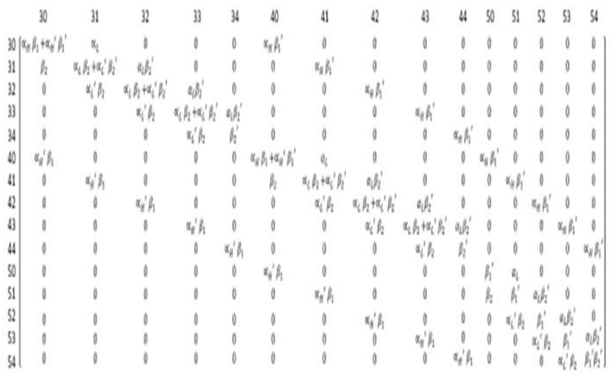

Figure.8. Finite Generator Matrix for system 2

The Boundary and Border Matrices (sub-matrices) for system 2 are shown in Figure.9.

$$B_{00} = \begin{bmatrix} \propto_H \beta_1 + \propto_H{}' \beta_1{}' & \propto_L & 0 & 0 & 0 \\ \beta_2 & \propto_L \beta_2 + \propto_L{}' \beta_2{}' & \alpha_L \beta_2{}' & 0 & 0 \\ 0 & \propto_L{}' \beta_2 & \propto_L \beta_2 + \propto_L{}' \beta_2{}' & \alpha_L \beta_2{}' & 0 \\ 0 & 0 & \propto_L{}' \beta_2 & \propto_L \beta_2 + \propto_L{}' \beta_2{}' & \alpha_L \beta_2{}' \\ 0 & 0 & 0 & \propto_L{}' \beta_2 & \beta_2{}' \end{bmatrix}$$

$$B_{01} = \begin{bmatrix} \propto_H \beta_1{}' & 0 & 0 & 0 & 0 \\ 0 & \propto_H \beta_1{}' & 0 & 0 & 0 \\ 0 & 0 & \propto_H \beta_1{}' & 0 & 0 \\ 0 & 0 & 0 & \propto_H \beta_1{}' & 0 \\ 0 & 0 & 0 & 0 & \propto_H \beta_1{}' \end{bmatrix}$$

$$B_{10} = \begin{bmatrix} \propto_H{}' \beta_1 & 0 & 0 & 0 & 0 \\ 0 & \propto_H{}' \beta_1 & 0 & 0 & 0 \\ 0 & 0 & \propto_H{}' \beta_1 & 0 & 0 \\ 0 & 0 & 0 & \propto_H{}' \beta_1 & 0 \\ 0 & 0 & 0 & 0 & \propto_H{}' \beta_1 \end{bmatrix}$$

Figure.9. Boundary and Border Matrices for system 2

The Repetitive Matrices (sub-matrices) for system 2 are shown in Figure.10.

$$A_0 = \begin{bmatrix} \propto_H \beta_1{}' & 0 & 0 \\ 0 & \propto_H \beta_1{}' & 0 \\ 0 & 0 & \propto_H \beta_1{}' \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad A_b = \begin{bmatrix} \beta_1{}' & \propto_L & 0 \\ \beta_2 & \beta_1{}' & \propto_L \beta_2 \\ 0 & \propto_L{}' \beta_2 & \beta_1{}' \\ 0 & 0 & \propto_L{}' \beta_2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} \propto_H \beta_1 + \propto_H{}' \beta_1{}' & \propto_L & 0 & 0 & 0 \\ \beta_2 & \propto_L \beta_2 + \propto_L{}' \beta_2{}' & \alpha_L \beta_2{}' & 0 & 0 \\ 0 & \propto_L{}' \beta_2 & \propto_L \beta_2 + \propto_L{}' \beta_2{}' & \alpha_L \beta_2{}' & 0 \\ 0 & 0 & \propto_L{}' \beta_2 & \propto_L \beta_2 + \propto_L{}' \beta_2{}' & \alpha_L \beta_2{}' \\ 0 & 0 & 0 & \propto_L{}' \beta_2 & \beta_2{}' \end{bmatrix}$$

$$A_2 = \begin{bmatrix} \propto_H \beta_1{}' & 0 & 0 & 0 & 0 \\ 0 & \propto_H \beta_1{}' & 0 & 0 & 0 \\ 0 & 0 & \propto_H \beta_1{}' & 0 & 0 \\ 0 & 0 & 0 & \propto_H \beta_1{}' & 0 \\ 0 & 0 & 0 & 0 & \propto_H \beta_1{}' \end{bmatrix}$$

Figure.10. The Repetitive Matrices (sub-matrices) for system 2

The Markov Chain for the system 2 is represented in Figure.11.

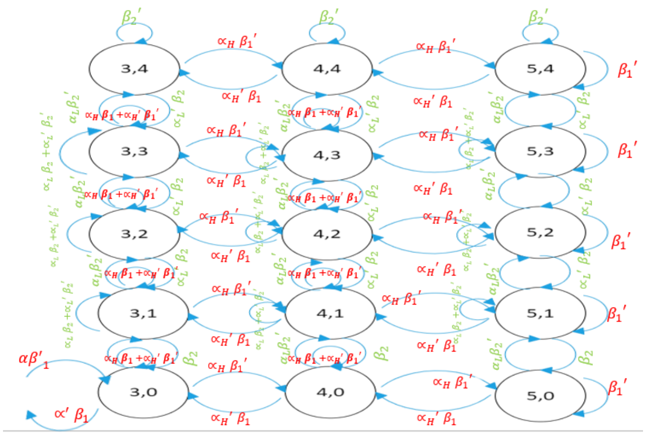

Figure.11. The Markov chain for the system 2

4. Results and Discussion

The analytical and simulation algorithms of the research are implemented in MATLAB. The arrival and services in the system both are discrete time functions and the system calculations are based on geometric distributions. The Figure.12 represents the number of low priority arrivals in queue 1 and 2 at different threshold values. In q1 and q2, a graph is produced between system usage and mean number. The maximum system utilization is 1 and the smallest system utilization is 0.

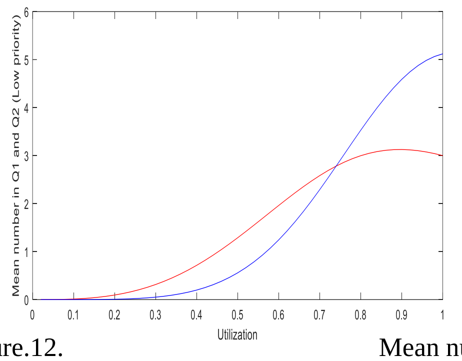

Figure.12. Mean number in the Queue1 and Queue2 (Low Priority).

The Figure.13 illustrates the number of high priority arrivals in queue 1 at different threshold values. This graph is shown for various service rates and system utilization for high priority customers.

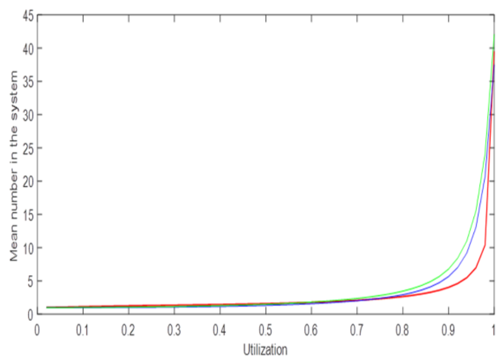

Figure.13. Mean number in the Queue1 (High Priority)

The Figure.14 shows the graphs between system utilization and average delay for different values of high priority arrivals ($\propto_H$), low priority arrivals ($\propto_L$), and servers (β).

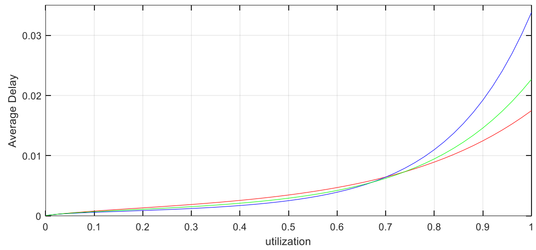

Figure.14. The average delay of the system

The Figure.15 displays the transient analysis of the system for different values of high and low priority arrivals. When the difference between S_T and S is small, the system exhibits few fluctuations and quickly approaches a steady state. However, as the difference grows, the system experiences more fluctuations.

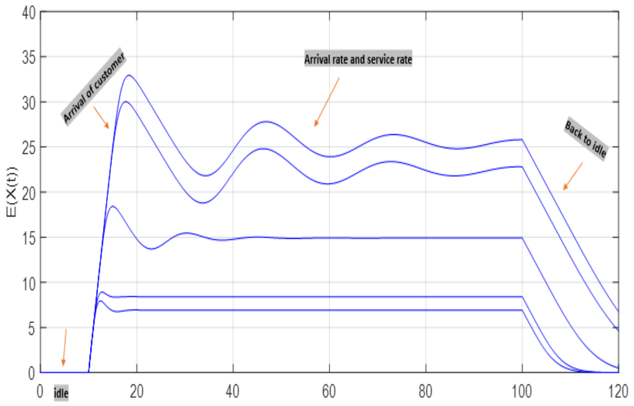

Figure.15. Transient Analysis for Various Values of Low and High Priority Arrivals

4. Conclusion

In this paper, performance analysis of threshold-based scheduling systems has been presented. The purpose of research is to avoid system congestion. Hence, a queuing mechanism based on internal and external criteria is being developed. Multidimension Markov chain and Matrix Geometric method are used as core techniques. Arrivals and services both are taken discrete in the system, based on geometric distribution. System size, delay response, and transient analysis for two types of arrivals in their respective systems are discovered and observed. Finally, the low and high priority arrivals are distinguished on the basis of threshold levels. The simulation and analytical results were completely identical. The queue's capacity and the service rate can be changed depending on the system's requirements.

References

[1] N. Madadi, A. H. Roudsari, K. Y. Wong, and M. R. Galankashi, "Modeling and Simulation of a Bank Queuing System," pp. 209–215, 2013, doi: 10.1109/CIMSim.2013.41.

[2] J. Walraevens, D. Fiems, and H. Bruneel, "Performance Analysis of Priority Queueing Systems in Discrete Time".

[3] E. G. Medvedeva, A. v. Gorbunova, Y. v. Gaidmaka, and K. E. Samouylov, "A Discrete Queueing Model for Performance Analysis of Scheduling Schemes in Multi-User MIMO Systems," *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops*, vol. 2019-Octob, 2019, doi: 10.1109/ICUMT48472.2019.8970876.

[4] E. Under and B. Constraint, "Scheduling Two Gauss – Markov Systems : f f ; f," vol. 60, no. 4, pp. 2038–2042, 2012.

[5] J. Arne, G. Selinka, and R. Stolletz, "Performance analysis of time-dependent queueing systems : Survey and classi fi cation $," vol. 63, pp. 170–189, 2016, doi: 10.1016/j.omega.2015.10.013.

[6] L. Guan *et al.*, "A Discrete-Time Performance Model for Congestion Control Based on Random Early Detection Using Queue Thresholds," vol. 2, pp. 1–6, doi: 10.5013/IJSSST.a.22.02.07.

[7] B. Ahmed and R. Ali, "A Discrete Event Simulation for the Analytical Modeling of M / D / 1 Queues : Output Buffer of an ATM Multiplexer," pp. 239–244, 2016.

[8] H. Wang, H. Huang, Y. Li, and Y. Yang, "Condition-Based Maintenance With Scheduling Threshold and Maintenance Threshold," pp. 1–12, 2015.

[9] V. Burger, T. Zinner, L. A. M. Dinh-xuan, and F. Wamser, "A Generic Approach to Video Buffer Modeling Using Discrete-Time Analysis," vol. 14, no. 2, 2018.

[10] D. E. Q. Alex S. Leong, Subhrakanti Dey, "Sensor, Transmission scheduling for remote state estimation and control with an energy harvesting Automatica," vol. Volume 201, p. Pages 0005-1098, 2018.

[11] D. J. Antunes and M. H. I. Balaghi, "Consistent Event-Triggered Control for Discrete-Time Linear Systems With Partial State Information," *IEEE Control Systems Letters*, vol. 4, no. 1, pp. 181–186, 2020, doi: 10.1109/LCSYS.2019.2922187.

[12] F. A. Nada, "UNFINISHED WORK & WAITING TIME OF GENERAL DISCRETE-TIME COMMUNICATION," *2020 International Conference on Innovative Trends in Communication and Computer Engineering (ITCE)*, pp. 122–125, doi: 10.1109/ITCE48509.2020.9047804.

[13] S. Agarwal, "Availability and maintenance modeling of a batch service queuing system," no. 1985, 2021, doi: 10.1108/IJQRM-12-2020-0407.

[14] O. Patricia, A. S. Alfa, and B. T. J. Maharaj, "Managing the harvested energy in wireless sensor networks : A priority Geo / Geo / 1 / k approach with threshold," *Energy Reports*, vol. 8, pp. 2448–2461, 2022, doi: 10.1016/j.egyr.2022.01.159.

[15] K. Jeganathan, N. Anbazhagan, S. Amutha, and V. Vinitha, "A queueing inventory system with threshold based priority service and a flexible server," 2020, doi: 10.1080/09720529.2020.1744811.

[16] S. Squillantet, C. H. Xiat, D. D. Yao, and L. Zhangt, "Threshold-Based Priority Policies for Parallel-Server Systems with Affinity Scheduling r," pp. 2992–2999, 2001.

About Authors

Engr. Mahnoor Memon

Master of Engineering Student,
Department of Electronic Engineering,
Mehran University of Engineering and Technology, Jamshoro.

Prof. Dr. Wajiha Shah

Professor,
Department of Electronic Engineering,
Mehran University of Engineering and Technology, Jamshoro

Engr. Qudsia Memon,

Lecturer,
Department of Electronic Engineering,
Mehran University of Engineering and Technology, Jamshoro