International Journal of Engineering, Science and Mathematics

Vol. 11 Issue 11, November 2022, ISSN: 2320-0294 Impact Factor: 6.765

Journal Homepage: http://www.ijesm.co.in, Email: editorijmie@gmail.com

Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Application of Queuing Theory for Improved Efficiency in Healthcare Systems

Balveer Saini,

Research Scholar,
Department of Mathematics, MS Brij University, Bharatpur (Raj.)

Dr.Kailash Chand Sharma

Principal(Retd.), Rajasthan Higher Education

Abstract

Queuing theory is a branch of operations research that deals with the study of queues, waiting lines, and the optimization of processes in various systems. In recent years, healthcare systems have faced numerous challenges in managing patient flow, resource allocation, and overall efficiency. This paper explores the application of queuing theory to address these challenges and improve the efficiency of healthcare systems. The paper begins by providing an overview of queuing theory, including key concepts and mathematical models commonly used. It then delves into the specific challenges faced by healthcare systems, such as long patient waiting times, overcrowded emergency departments, and inefficient resource allocation. By applying queuing theory, healthcare administrators and policymakers can gain valuable insights into the factors contributing to these challenges and develop data-driven solutions. One of the important application of queuing theory in healthcare is predicting patient flow and waiting times. By modelling on patient arrivals, service times, and staff availability, healthcare facilities can optimize their scheduling and staffing to reduce waiting times and improve patient satisfaction. Additionally, queuing theory can be used to analyse the impact of implementing various strategies, such as appointment systems, triage protocols, and resource reallocation.

Keywords: Queueing Theory, Healthcare System, Emergency Department etc.

Introduction

Healthcare systems play a vital role in ensuring the well-being of individuals and communities by providing essential medical services. However, the ever-increasing demand for healthcare services, coupled with limited resources, has put immense pressure on healthcare facilities worldwide. Long patient waiting times, overcrowded emergency departments, and inefficient resource allocation are just a few of the challenges that healthcare administrators and policymakers face in their quest for optimal system performance.

To address these challenges, various disciplines within operations research and management science have been applied to healthcare management. One such powerful tool is queuing theory, a branch of operations research that focuses on the study of queues, waiting lines, and the mathematical modelling of processes. Queuing theory has proven to

be highly applicable in diverse industries, from telecommunications to manufacturing, and its potential to revolutionize healthcare systems is becoming increasingly recognized.

This paper explores the application of queuing theory in healthcare systems to achieve improved efficiency, enhanced patient outcomes, and overall system optimization. By understanding the principles and methodologies of queuing theory, healthcare administrators can gain valuable insights into the factors affecting patient flow and resource utilization. Leveraging this knowledge, they can implement data-driven strategies to minimize waiting times, increase the effectiveness of medical services, and maximize the utilization of available resources.[1]

The paper examinevarious real-world applications of queuing theory in healthcare, such as predicting patient flow, optimizing emergency department operations, and streamlining patient admissions. Additionally, it will explore how queuing theory can be utilized to improve healthcare supply chain management, ensuring the availability of critical medical resources when and where they are needed. It is essential to recognize that implementing queuing theory in healthcare systems requires a multidisciplinary approach, involving collaboration between healthcare professionals, operations researchers, and policymakers. The synergy between domain expertise and analytical tools can lead to evidence-based decision-making, ultimately benefiting patients, healthcare providers, and the overall healthcare ecosystem.

Application of Queuing Theory

The application of queuing theory in healthcare systems offers valuable insights and optimization opportunities to enhance patient care, resource management, and overall system efficiency. Some of the key applications include:

Patient Flow Management: Queuing theory helps healthcare facilities manage patient flow by analysing arrival patterns and service times. This enables administrators to design efficient patient pathways, reducing waiting times and ensuring timely access to care.

Resource Allocation: By modelling queueing systems, healthcare units can optimize resource allocation, such as staff scheduling and equipment utilization. This leads to improved resource efficiency and cost-effectiveness.

Emergency Department Triage: Queuing models assist in optimizing emergency department triage processes, prioritizing patients based on severity and ensuring prompt medical attention for critical cases.

Operating Room Scheduling: Queuing theory is applied to optimize operating room schedules, minimizing idle time and maximizing utilization while ensuring smooth surgical procedures.

Pharmacy Services: In pharmacy settings, queuing theory helps to manage medication dispensing queues, reducing waiting times for patients to receive their prescriptions.

Appointment Scheduling: Healthcare facilities use queuing theory to optimize appointment scheduling, reducing patient wait times for consultations and treatments.

Capacity Planning: Queuing models aid in capacity planning, allowing healthcare systems to anticipate future demands and ensure sufficient resources are available to meet patient needs.

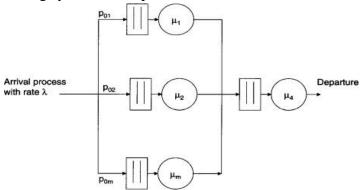
Predictive Analytics: Queuing theory's mathematical models facilitate predictive analytics, helping healthcare administrators forecast patient volumes and plan accordingly. By leveraging queuing theory in healthcare systems, hospitals and clinics can improve patient experience, increase operational efficiency, and ultimately deliver better quality care to patients. It enables data-driven decision-making, helping healthcare facilities address challenges and enhance their services for the benefit of both patients and staff.[2]

Need of the Study

The need for the study on the application of queuing theory in healthcare systems stems from the growing challenges faced by healthcare facilities in managing patient flow, resource allocation, and overall efficiency. Long waiting times, overcrowded emergency departments, and inefficient utilization of medical resources adversely affect patient outcomes and satisfaction. Queuing theory offers a systematic and data-driven approach to address these issues by providing insights into the dynamics of patient flow, waiting times, and service utilization. By understanding the underlying principles of queuing theory, healthcare administrators can make informed decisions to optimize resource allocation, enhance staff scheduling, and streamline patient admissions. Implementing queuing models can lead to reduced waiting times, increased patient throughput, and improved utilization of healthcare resources, ultimately resulting in better patient care and overall system efficiency. The study aims to bridge the gap between theory and practice, empowering healthcare professionals with tools to design efficient healthcare processes, respond effectively to fluctuations in patient demand, and create patient-centric systems that deliver timely and high-quality medical services.

Queueing network model

A queueing network model is a powerful mathematical framework used to analyse and optimize complex systems with interconnected queues. It is particularly applicable to scenarios where entities (e.g., customers, patients, or tasks) move through multiple service stations in a predefined sequence. In this model, each service station represents a queue, and the entities follow predefined routes or paths as they progress through the network. Queueing network models find widespread use in various fields, including computer networks, manufacturing systems, transportation, and healthcare.



By using mathematical techniques, such as Markov chains and differential equations, these models can predict system performance metrics like throughput, response times, and resource utilization. The flexibility of queueing network models allows for the analysis of different system configurations and the evaluation of various policy changes before

implementation. This enables decision-makers to optimize resource allocation, identify potential bottlenecks, and enhance overall system efficiency. queueing network models provide valuable insights into the dynamics of interconnected queues, helping businesses and organizations make informed decisions to improve their processes and services, resulting in better performance, reduced costs, and increased customer satisfaction.[3]

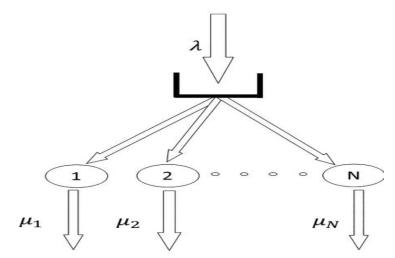


Figure 1. Queueing system with heterogeneous servers.

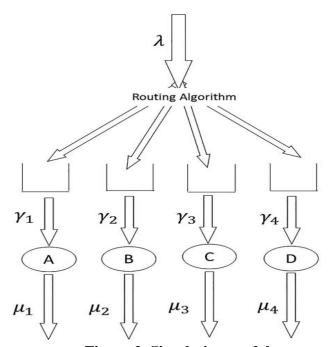


Figure 2. Simulation model.

The M/M/S model in queuing theory can be defined using the following mathematical terms:

- λ (lambda): The arrival rate of entities (customers, patients, etc.) to the system, following a Poisson distribution. It represents the average number of arrivals per unit of time.
- μ (mu): The service rate of each server, also following an exponential distribution. It represents the average number of entities that a single server can serve in one unit of time.
- S: The number of servers in the system, indicating the total capacity to serve entities simultaneously.

 ρ (rho): The utilization factor of the servers, calculated as $\rho = \lambda / (S * \mu)$. It represents the proportion of time the servers are busy.

L (average number of entities in the system): $L = (\lambda * W)$, where W is the average time an entity spends in the system.

W (average time an entity spends in the system): $W = 1 / (\mu - \lambda)$, representing the average time an entity spends in the system, including both waiting and service times.

 L_q (average number of entities in the queue): $L_q = (\lambda^2) / (\mu * (\mu - \lambda))$, representing the average number of entities waiting in the queue.

 W_q (average time an entity spends waiting in the queue): $W_q = \lambda / (S * (\mu - \lambda))$, representing the average time an entity spends waiting in the queue.

These mathematical terms enable the calculation of important performance measures in the M/M/S model, helping to analyse and optimize the efficiency of queuing systems with multiple servers in various applications, including healthcare.[4].

Literature Review

Fomundam, S., et al (2007)Queuing theory, a branch of operations research, has found widespread applications in the healthcare sector due to its ability to analyse and optimize the flow of patients in various medical facilities. This abstract provides a concise overview of queuing theory applications in healthcare. Healthcare organizations are often faced with the challenge of managing patient flow efficiently to reduce waiting times, enhance resource utilization, and improve overall service quality. Queuing theory offers valuable insights into understanding patient arrival patterns, service times, and queue lengths, allowing healthcare providers to make informed decisions. Applications of queuing theory in healthcare encompass a broad range of areas, such as emergency departments, outpatient clinics, operating rooms, and pharmacy services. By modeling patient flow, queuing theory aids in identifying bottlenecks, predicting waiting times, and optimizing staff scheduling and resource allocation. Moreover, it assists in designing effective triage systems and estimating capacity requirements to accommodate patient demand effectively.[1]

Peter, P. O., &Sivasamy, R. (2019).Queueing theory techniques have emerged as a valuable tool with real-world applications in healthcare systems. This abstracthighlights the significance of queuing theory in improving healthcare services. By analysing patient arrival patterns, service times, and queue lengths, queuing theory provides valuable insights for optimizing resource allocation and reducing patient waiting times. Its applications span across various healthcare settings, including emergency departments, outpatient clinics, operating rooms, and pharmacy services. Through the use of queuing models, healthcare administrators can make informed decisions about staff scheduling, capacity planning, and facility design, resulting in enhanced system efficiency and patient flow. queuing theory aids in identifying bottlenecks in healthcare processes, enabling targeted improvements to enhance overall service quality. By integrating queuing theory techniques, healthcare systems can better manage patient demand, streamline operations, and allocate resources more effectively.[2]

Lakshmi, C., & Iyer, S. A. (2013). This literature review explores the application of queueing theory in the field of health care. Queuing theory has gained significance as a powerful analytical tool in health care systems to optimize patient flow and resource utilization. By analysing patient arrivals, service times, and queue lengths, queuing models

assist in identifying areas for improvement, such as reducing waiting times and increasing staff efficiency. The review highlights the diverse applications of queuing theory in different health care settings, including emergency departments, outpatient clinics, and hospitals, as well as its relevance in capacity planning and resource allocation. Additionally, it examines how queuing theory aids in enhancing patient experience, streamlining processes, and ultimately contributing to better health care delivery. This abstract underscores the valuable insights that queuing theory offers to the health care industry.[3]

Green, L. (2006). Queueing analysis, based on queuing theory principles, has emerged as a valuable tool to improve healthcare processes. By studying patient arrivals, service times, and queue lengths, queueing analysis provides valuable insights into the dynamics of patient flow through various healthcare services, including emergency departments, outpatient clinics, and operating rooms. Through the application of mathematical models, healthcare administrators can identify bottlenecks, predict waiting times, and optimize resource allocation to enhance operational efficiency. This abstract underscores the practical importance of queueing analysis in healthcare, offering a systematic approach to streamline processes, reduce waiting times, and improve overall patient experience. By integrating queueing analysis, healthcare systems can make data-driven decisions to deliver more effective and patient-centric care.[4]

Kalwar, M. A., et al (2021)Queuing theory and discrete event simulation have found diverse applications in healthcare units, providing valuable insights and optimization opportunities to improve patient care and resource management. Queuing theory is utilized to analyze patient flow, waiting times, and service utilization, enabling healthcare units to identify inefficiencies and bottlenecks in processes such as appointment scheduling, emergency room triage, and pharmacy services. By modeling these systems, queuing theory helps to predict patient wait times and optimize resource allocation for better overall performance. Discrete event simulation complements queuing theory by simulating detailed patient interactions and resource usage. This approach allows healthcare units to test various scenarios, assess the impact of process changes, and evaluate potential improvements before implementation. Discrete event simulation is particularly useful in complex healthcare settings like hospitals, where interactions between multiple departments and resources play a crucial role. The combination of queuing theory and discrete event simulation empowers healthcare units to make informed decisions, enhance patient outcomes, and optimize resource utilization, ultimately leading to more efficient and effective healthcare delivery.[5]

Queuing theory and health care

Queuing theory plays a crucial role in health care systems, utilizing mathematical models to analyse patient flow, waiting times, and resource utilization. In health care, patients arrive at a facility and enter a queue for service, such as consultations, tests, or treatments. These queues can lead to delays and affect patient satisfaction and overall system efficiency.[6]

Mathematical equations, based on queuing theory principles, allow healthcare administrators to calculate important performance measures, such as the average waiting

time, the average number of patients in the queue, and the utilization of healthcare resources.

One commonly used queuing model in healthcare is the M/M/1 queue, where "M" stands for Poisson arrivals (random arrival rate), "M" stands for exponential service times (random service rate), and "1" refers to a single server. By analysing this model, healthcare units can determine the expected waiting times and the probability that a patient has to wait for service. Let P_n be the probability of system and n units are present in the system in time $t+\Delta t$, i.e. the system is in E_n state.

Then we can find the probability that the system is in state E_n at timet+ Δt by the given formulae

And the study state equations for M/M/1 model are

$$\frac{dP_n}{dt} = \lambda P_{n-1} - (\lambda + \mu) P_n + \mu P_{n+1}$$

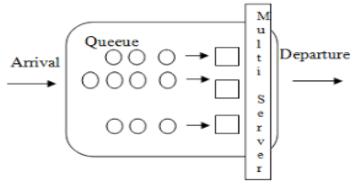
$$\frac{dP_0}{dt} = -\lambda P_0 + \mu P_1$$
(2)

Additionally, more complex queuing networks can be applied to healthcare systems, considering multiple service stations and patient routes. Mathematical techniques, such as Markov chains and differential equations, help in modelling and analysing these intricate networks.[7]

By employing queuing theory with mathematical precision, healthcare facilities can optimize patient flow, reduce waiting times, and allocate resources efficiently, ultimately leading to improved patient care and enhanced healthcare delivery.

The M/M/S Model

The M/M/S model is a widely used queuing model that is applicable to situations where entities, such as customers or patients, arrive at a service system with a Poisson arrival rate (random arrival rate) and are served by multiple identical servers (S) with exponential service times (random service rate). The model is characterized by the notation M/M/S, where "M" stands for Poisson arrivals, "M" denotes exponential service times, and "S" represents the number of servers.[8]



In the M/M/S model, the arrival process follows a Poisson distribution, meaning that the time between consecutive arrivals is exponentially distributed. Similarly, the service times are exponentially distributed, representing the time each server takes to serve a single entity.

Key performance measures that can be calculated using the M/M/S model including the average number of entities in the system (L), the average time an entity spends in the system (W), the utilization of the servers (ρ), and the probability of the system being in a particular state (P_n).

Let E_n be the state of system i.e. n units are present in the system in time $t+\Delta t$ and $n \geq S$.

Then we can find the probability for the Queueing system M/M/S at timet+ Δt by the given formula

$$P_{n}(t+\Delta t) = P_{n}(t)[1-(\lambda+S\mu)\Delta t] + P_{n-1}(t)[\lambda\Delta t + o(\Delta t)^{2}] + P_{n+1}(t)[(S\mu\Delta t + o(\Delta t)^{2}]....(4)$$

And the study state equations for M/M/S model are

Measuring the performance of a queuing system involves various performance metrics that can be expressed using mathematical equations. Here are some common performance metrics and their corresponding equations:

Utilization (ρ): The utilization of the system represents the fraction of time the server is busy. It's the ratio of the average service rate (μ) to the arrival rate (λ) of customers.

Equation: $\rho = \lambda / \mu$

Throughput (X): Throughput is the average number of customers that the system can process per unit of time.

Equation: $X = \lambda * P(system)$

Here, P(system) is the probability of the system having at least one customer.

Average Queue Length (L_q): The average number of customers waiting in the queue.

Equation: $L_q = (\lambda^2) / (\mu * (\mu - \lambda))$

Average Time in Queue (W_q) : The average time a customer spends waiting in the queue.

Equation: $W_q = L_q / \lambda$

Average Time in the System (W): The average time a customer spends in the entire system (waiting time + service time).

Equation: $W = W_q + (1 / \mu)$

Probability of zero Patients in the System (P(0)): The probability of having no customers in the system.

Equation: $P(0) = 1 - \rho$

Probability of N Patients in the System (P(N)): The probability of having exactly N customers in the system.

Equation: $P(N) = ((1 - \rho) * (\rho^{N}))$

Probability of Waiting (P_w): The probability that a customer has to wait in the queue before being served.

Equation: $Pw = \lambda / \mu$

Little's Law: Little's Law relates the average number of customers in the system (L) to the average time a customer spends in the system (W) and the arrival rate (λ).

Equation: $L = \lambda * W$

These equations are fundamental equations in analysing the performance of queuing systems and can be used to evaluate and optimize the system's efficiency and customer experience. Keep in mind that the specific performance metrics and equations might vary depending on the queuing system's characteristics and the type of queue (e.g., M/M/1, M/M/c, etc.).

The M/M/S model is valuable in various applications, including healthcare systems with multiple parallel servers, such as multiple doctors in a clinic or multiple operating rooms in a hospital. By utilizing this queuing model, healthcare administrators can optimize resource allocation, predict waiting times, and improve overall system efficiency to enhance patient care and satisfaction.[9]

$$P(W_q > 0) = \sum_{l=0}^{c} \pi_{N+l}$$

$$= \pi_0 \frac{\lambda^N}{N! \prod_{i=1}^{N} \mu_i^{N_i}} \sum_{l=0}^{c} \frac{\prod_{i=0}^{l-1} \lambda_i}{\prod_{i=1}^{l} \mu_i}$$

where π is given in (4) above and we can obtain the queue length distribution as

$$P(L_{q} = l | W_{q} > 0) = \frac{\pi_{N+l}}{\sum_{i=0}^{c} \pi_{N+i}}$$

$$= \frac{\prod_{i=0}^{l-1} \lambda_{i}}{\prod_{i=1}^{l} \mu_{i}},$$

$$\sum_{k=0}^{c} \frac{\prod_{i=0}^{k-1} \lambda_{i}}{\prod_{i=1}^{k} \mu_{i}},$$

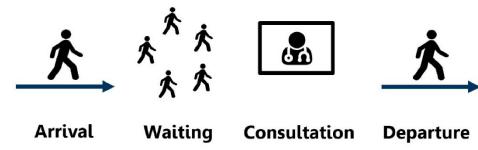
$$\forall l = 0, 1, \dots, c.$$

$$\begin{split} P(L_q = l) &= P(L_q = l | W_q > 0) P(W_q > 0) \\ &= \frac{\prod_{i=0}^{l-1} \lambda_i}{\prod_{i=1}^{l} \mu_i} P(W_q > 0), \\ &= \frac{\sum_{k=0}^{c} \frac{\prod_{i=0}^{k-1} \lambda_i}{\prod_{i=1}^{k} \mu_i}}{V(W_q > 0)}, \\ \forall \ l = 1, 2, \dots, c. \end{split}$$

The primary performance measure of interest in the loss model is the steady-state loss probability, denoted as P(N). This probability represents the likelihood that the system is in a state where all the servers are busy, i.e., all the service channels are occupied, and an incoming customer cannot immediately be served. In this context, assuming Poisson arrivals, the steady-state loss probability (P(N)) can be expressed as the probability that an arriving customer is blocked (denoted as P(block)).

Average waiting time for OPD section with one extra doctor

To calculate the average waiting time for the OPD (Outpatient Department) section with one extra doctor, we can use the M/M/S queuing model, where "M" represents Poisson arrivals (patient arrival rate), "M" represents exponential service times (average time taken by a doctor to see a patient), and "S" is the number of servers (doctors).



Let's assume the current number of doctors in the OPD section is "S" (before adding the extra doctor), and the arrival rate of patients is " λ " (average number of patients arriving per unit of time). The current average service time of a doctor is " $1/\mu$ " (where μ is the service rate of a single doctor).

The utilization factor of the current system is given by $\rho = \lambda / (S * \mu)$, representing the proportion of time the current doctors are busy.

After adding one extra doctor, the new number of servers will be "S + 1". So, the new utilization factor will be $\rho' = \lambda / ((S + 1) * \mu)$.

Now, the average waiting time for the OPD section can be calculated as:

Average Waiting Time = $(\rho / (S * \mu)) / (1 - \rho)$ [for the current system]

Average Waiting Time with one extra doctor = $(\rho' / ((S + 1) * \mu)) / (1 - \rho')$ [for the system with one extra doctor]

By comparing the average waiting times for the current system and the system with one extra doctor, we can assess the impact of adding the extra doctor on reducing patient waiting times in the OPD section.[10]

Problem Statement

The problem statement focuses on utilizing Queuing Theory to enhance efficiency in healthcare systems. Healthcare facilities often face challenges in managing patient flow and resource utilization, resulting in long waiting times, inefficient processes, and decreased patient satisfaction. The aim of this study is to apply Queuing Theory principles to analyse patient arrivals, service times, and queue lengths within healthcare units. By doing so, the study seeks to identify bottlenecks and inefficiencies, predict waiting times, and optimize resource allocation for better overall performance.[11]

The research will utilize mathematical models, such as the M/M/S model, to calculate crucial performance metrics like average waiting times, utilization of healthcare resources, and queue lengths. By analysing these metrics, healthcare administrators can make data-driven decisions to streamline patient flow, reduce waiting times, and allocate resources more effectively. Ultimately, the study aims to contribute to improved patient experience, enhanced healthcare delivery, and increased operational efficiency in healthcare systems.[12]

Conclusion

In conclusion, the application of queuing theory in healthcare systems has demonstrated its potential to significantly improve efficiency and patient care. By analysing patient flow, service times, and queue lengths, queuing theory helps healthcare administrators identify areas for optimization and make informed decisions to streamline operations. This leads to reduced waiting times, better resource allocation, and enhanced overall service quality. Queuing theory has proven particularly valuable in emergency departments, outpatient clinics, and operating rooms, where timely patient care is critical. Through its mathematical models, queuing theory allows healthcare units to predict patient waiting times and plan resources accordingly, leading to more effective patient management, when combined with discrete event simulation, queuing theory enables healthcare units to simulate complex interactions between various departments and resources, helping them test different scenarios and optimize processes before implementation. By leveraging the insights provided by queuing theory, healthcare systems can create more patient-centric, efficient, and well-organized environments, ultimately leading to improved healthcare outcomes and higher patient satisfaction. As the healthcare landscape continues to evolve, the continued application of queuing theory holds the promise of driving further advancements in healthcare delivery and resource management.

References

- [1] Fomundam, S., & Herrmann, J. W. (2007). A survey of queuing theory applications in healthcare.
- [2] Peter, P. O., &Sivasamy, R. (2019). Queueing theory techniques and its real applications to health care systems—Outpatient visits. *International Journal of Healthcare Management*.
- [3] Lakshmi, C., & Iyer, S. A. (2013). Application of queueing theory in health care: A literature review. *Operations research for health care*, 2(1-2), 25-39.
- [4] Green, L. (2006). Queueing analysis in healthcare. *Patient flow: reducing delay in healthcare delivery*, 281-307.
- [5] Green, L. V., Soares, J., Giglio, J. F., & Green, R. A. (2006). Using queueing theory to increase the effectiveness of emergency department provider staffing. *Academic Emergency Medicine*, *13*(1), 61-68.
- [6] Kalwar, M. A., Marri, H. B., Khan, M. A., &Khaskheli, S. A. (2021). Applications of queuing theory and discrete event simulation in health care units of Pakistan. *International Journal of Science and Engineering Investigations*, 10(109), 6-18.
- [7] Alhaag, M. H., Aziz, T., &Alharkan, I. M. (2015, March). A queuing model for health care pharmacy using software Arena. In 2015 International Conference on Industrial Engineering and Operations Management (IEOM) (pp. 1-11). IEEE.
- [8] Adeleke, R.A., Ogunwale, O.D., Halid O.Y. (2009) An application of queueing theory to waiting time of out-patients in hospitals. Pacific Journal of Science and Technology, 10(2), 270-274.
- [9] Au-Yeung, S.W.M., Harrison, P.G., Knottenbelt, W.J. (2007) A Queueing Network Model of Patient Flow in an Accident and Emergency Department, Valuetools 07, October 23-25, Nantes, France.

- [10] Biju M. K, Naeema, K., Faisal U (2011) Application of Queueing Theory in Human Resource Management in Health Care, ICOQM-10, June 28-30, 1019-1027.
- [11] De Bruin, A.M., Rossum, A.C., Visser, M.C., Koole, G.M (2007) Modelling the emergency cardiac in-patient flow: an application of queueing theory, Health Care Management Science, 10, 125-137.
- [12] Chaussalet T.J., Xie H., Millard P (2006) A Closed Queueing Network Approach to the Analysis of Patient Flow in Health Care Systems, Methods Information Medicine, 5.