IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, April 2026
Impact Factor: 8.2

Analysis of Idle Time in M/M/1 Queueing System with Arrival and Service Rate Variations

Dr Naveen Kumar[1] and Sarla[2]
Professor, Department of Mathematics, Baba Mastnath University, Asthal Bohar, Rohtak[1]
Research Scholar, Department of Mathematics, Baba Mastnath University, Asthal Bohar, Rohtak[2]
naveenkapilrtk@gmail.com and sarla.kaushik21@gmail.com

Abstract: *This paper examines the notion of idle time within an M/M/1 queueing model, through the analysis of the correlation between service rate (μ) and its arrival rate (lambda). The idle time, which is the time when the server is not occupied, is important in terms of efficiency in the system and resources. The study uses probabilistic modeling with Poisson arrival and exponential service distributions in order to obtain theoretical forms of utilization factor and idle time.*
In addition, the research paper will evaluate the effect of the changes in λ and μ on the system performance indicators, including server utilization, waiting time and idle probability. The comparison of theoretical outcomes with simulated data is done to authenticate the model. The trends and relations among parameters are noted on the graphical representation. Results indicate that idle time reduces as the rate of arrival rises and it grows as the service rates rise.
The paper offers useful information to streamline service systems in practical use of call center, health facilities, and manufacturing where there is a need to balance efficiency and cost..

Keywords: Queueing Theory, Idle Time, M/M/1 Model, Arrival Rate, Service Rate, Utilization Factor

I. INTRODUCTION

Queueing theory is a major aspect of applied mathematics and operations research, which is concerned with the investigation of waiting lines or queues. It offers a methodical structure to the analysis of systems whereby a customer, a data packet or even a job, come to be served by one or more service facilities. The field of queuing theory has over the years become very crucial in various industries such as telecommunications, health care, transport, manufacturing and computer networks where optimization of resources and services is very important.
The M/M/1 queueing system is one of the most common and most discussed models of the many types of queueing models since it is analytically tractable and of practical relevance. The arrivals in this model are Poisson process implying that there is no correlation between the inter-arrival times and that these times are exponentially distributed. In the same way, service time is also distributed exponentially and the system has only one server which serves the customers in a first come, first serve (FCFS) format. The M/M/1 model has been shown to offer a good understanding of how more complex queueing systems operate even though it is a very simple model.
Idle time is one of the most important key performance measures of queueing systems because it is the time that the server is not busy as a result of no customers in the system. Idle time is a key variable that determines service system efficiency and effectiveness. It is directly proportional to the utilization factor, which is used to determine the time spent by the server being busy. The idle time to server utilization ratio is crucial since it shows the current level of system resource utilization.
Theoretically, idle time is directly associated with stability and performance indicators of a system (average waiting time, queue length, and throughput). Specifically, the correlation between the service rate (μ) and the arrival rate (λ) will help to define whether the system is operating at high efficiency. When the arrival rate is much smaller than the service rate, the server has more idle time meaning that the available resources are not fully utilized. On the other hand,

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, April 2026 Impact Factor: 8.2

as the arrival rate is close to the service rate, the system is highly utilized and thus, idle time is minimized at the expense of congestion and waiting time is increased.

Practically, the idle time is very important to be handled in real-life scenarios. Unutilized idle time may lead to resource wastage and higher costs of operations particularly in systems like call centers, hospitals and manufacturing units. Conversely, under-utilizing idle time could result in over-utilized systems, lower quality of service, and unhappy customers. Thus, one of the goals of service system design and administration is to achieve an ideal compromise between system underuse and overuse.

This paper will be able to study idle time in an M/M/1 queueing model by exploring the effects of the change in the arrival rate (λ) and service rate (μ). The study aims at giving a deeper insight into the effects of system parameters on performance by looking at both the theoretical associations and real implications of the research. The results of the research will help in creating effective and affordable service systems in different fields.

II. LITERATURE REVIEW

The idle time in the queueing systems has received much attention in the recent years because of its effects on the efficiency of the system and its resources. Various researchers have studied various queueing models and service mechanism in order to learn the dynamics of idle time and its optimization.

Madabi et al. (2023) studied Erlang loss systems whose idle service discipline is the shortest. They concentrated their work on the maximum use of servers, introducing maintenance measures like the inversion of stacks and replacement of servers. To study the behavior of the system and reduce the cost of its operations in the long-term, the researchers created a continuous-time Markov chain model. Their results note the need to have good idle server management, which correlates directly to the reduction of idle time in the system and the performance of the system[1].

Azhagappan and Deepa (2020) involved exploring a Markovian queueing model with the policies of customer impatience, balking, and working vacation. Their model also examined the exit of customers because of the waiting delays and the impact of idle time on the system dynamics that the booking phase is in. The research has given both temporary and numerical examples and demonstrates that idle time could be employed wisely even during working holidays to ensure that the system is stable and responsive[2].

Liu and Liu (2021) researched the uncertain queueing system to examine the idle time and the waiting time in terms of the uncertainty theory. Their study formulated the analytical equations in estimating the idle time distribution when there are unpredictable arrival and service time. It was pointed out by the study that variability of parameters in the systems is a major factor that can determine the idle time and thus the need to ensure that the modeling adopted in real life applications should be flexible[3].

Ayyappan et al. (2020) examined a non-preemptive priority queueing system without preemptive retrial in which servers could break down, working vacation, and reneging were allowed. Their model also included realistic living conditions and they also studied system stability with the supplementary variable technique. Results showed that the breakdowns and retrials of the servers have a great influence on the idle time and system efficiency, and it is necessary to design a robust system[4].

---

[1] Madadi, M., Heydari, M. H., Maillart, L., Cassady, R., & Zhang, S. (2023). Erlang loss systems with shortest idle server first service discipline: Maintenance considerations. *IISE Transactions, 55*(10), 1008–1021.

[2] Azhagappan, A., & Deepa, T. (2020). Variant impatient behavior of a Markovian queue with balking reserved idle time and working vacation. *RAIRO-Operations Research, 54*(3), 783–793.

[3] Liu, Y., & Liu, B. (2021). Waiting time and idle time of uncertain queueing systems. *International Journal of General Systems, 50*(8), 871–890.

[4] Ayyappan, G., Udayageetha, J., & Somasundaram, B. (2020). Analysis of non-preemptive priority retrial queueing system. *International Journal of Mathematics in Operational Research, 16*(4), 480–498.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, April 2026 Impact Factor: 8.2

Zhong et al. (2022) to investigate the optimal idling strategies in many-server systems of queueing. Their analysis has revealed a non-idling policy is not necessarily optimal and that managed idle time might be used to achieve a balanced operation cost and service efficiency. The study presented theoretical information on how idle time can be strategically used instead of reducing it to a minimal level[5].

2.1 Research Gap

Although much has been done in the researches about queueing systems, there have been fewer studies that specifically concentrate on the correlation between idle time and the rate of varying arrival and service rates in the simple M/M/1 model. This paper will help address this gap by offering a comprehensive analytical and graphical study of the behavior of idle time.

III. OBJECTIVES OF THE STUDY

To examine wasted time in M/M/1 queueing system.
To investigate the correlation of the arrival rate (λ) and idle time.
To investigate the impact of service rate (μ) on the system performance.
To compare theoretical and simulated results.
To give graphical explanation of system behavior[6].

IV. MATHEMATICAL MODEL

The M/M/1 queueing system has a mathematical formulation that offers a systematic model of the analysis of idle time and system performance. The following section is the base equations, relation and assumption of the study.
Core Formula (Utilization Factor)

$$\boldsymbol{\rho = \frac{\lambda}{\mu}}$$

The utilization factor (ρ) represents the proportion of time the server is busy. It is defined as the ratio of arrival rate (λ) to service rate (μ).
When the value of rho is larger, it means that the utilization of the server is high however when the value of rho is smaller then it means that the server is not fully utilized. The efficiency of the utilization is a key parameter in the behavior of idle times and efficiency of the system[7].

4.1 Key Relations

Probability that System is Empty (Idle Time)

$$P_0 = 1 - \rho$$

The likelihood of the system having zero customers is indicated as $P0$, and this is also the idle probability of the server. This means that the server is not used when there are no people in the system. Thus, the idleness is directly connected to the complement of the utilization factor.
Condition for Stability

$$\lambda < \mu$$

For the M/M/1 queueing system to reach a steady-state condition, the arrival rate must be less than the service rate. If this condition is violated, the queue will grow indefinitely, leading to system instability.

[5] Zhong, Y., Ward, A. R., & Puha, A. L. (2022). Asymptotically optimal idling in the GI/GI/N+GI queue. *Operations Research Letters, 50*(3), 362–369.
[6] Erlang, A. K. (1909). On the theory of probabilities in telephone traffic and congestion analysis. Copenhagen: Nyt Tidsskrift for Matematik.
[7] Kendall, D. G. (1953). Stochastic processes and their applications in queueing theory. Journal of the Royal Statistical Society: Series B (Methodological), 15(2), 151–185.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, April 2026 Impact Factor: 8.2

This condition ensures that the server can handle incoming customers efficiently without excessive accumulation in the queue.

Idle Time Interpretation

$$Idle\ Time = 1 - \frac{\lambda}{\mu}$$

Idle time is mathematically expressed as the complement of the utilization factor. This equation clearly shows that idle time depends on the relative values of λ and μ.

When λ is small compared to μ, idle time is high

When λ approaches μ, idle time decreases

When λ is very close to μ, idle time approaches zero

Thus, idle time provides insight into how effectively system resources are utilized.

4.2 Additional Performance Measures

Although the primary focus of this study is idle time, other important performance measures in the M/M/1 queueing model are also useful for evaluating overall system behavior[8].

Average number of customers in the system:

$$L = \frac{\lambda}{\mu - \lambda}$$

This formula gives the expected number of customers present in the system, including both those waiting in the queue and the one being served. As the arrival rate approaches the service rate, the value of LLL increases rapidly, indicating congestion in the system.

Average number of customers in the queue

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

This represents the average number of customers waiting in the queue before receiving service. It helps in understanding the extent of crowding in the waiting line.

Average waiting time in the system

$$W = \frac{1}{\mu - \lambda}$$

This formula gives the expected total time a customer spends in the system, including both waiting time and service time.

Average waiting time in the queue

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

This represents the average time a customer spends waiting in the queue before being served.

These measures are interconnected and provide a broader understanding of queue performance. While idle time reflects server underutilization, the values of L, Lq, W, and Wq indicate congestion and customer delay. Together, they help in assessing whether the system is operating efficiently and whether the chosen arrival and service rates are balanced[9].

[8] Gross, D., Shortle, J. F., Thompson, J. M., & Harris, C. M. (1998). Fundamentals of queueing theory (3rd ed.). New York, NY: Wiley.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, April 2026 Impact Factor: 8.2

4.3 Assumptions of the Model

The mathematical model is based on the following standard assumptions:
Poisson Arrival Process: Customers arrive randomly and independently at a constant average rate (λ).
Exponential Service Time: Service times follow an exponential distribution with rate (μ).
Single Server: Only one server is available to serve customers[10].
Infinite Queue Capacity: There is no limit on the number of customers that can wait in the queue.
First-Come, First-Served (FCFS): Customers are served in the order of their arrival[11].

V. METHODOLOGY

The present study adopts a quantitative, analytical, and simulation-based approach to evaluate idle time in an M/M/1 queueing system. The methodology is designed to integrate theoretical modeling with computational validation in order to provide a comprehensive understanding of system behavior under varying conditions.

At the theoretical level, the study is based on the standard assumptions of the M/M/1 queueing model, where arrivals follow a Poisson distribution and service times are exponentially distributed. A single-server system with infinite queue capacity and a First-Come, First-Served (FCFS) discipline is considered. Using fundamental principles of probability theory and queueing analysis, key performance measures such as utilization factor, idle time, waiting time, and system stability are derived. Idle time is specifically calculated as the complement of the utilization factor, which depends on the ratio of arrival rate (λ) to service rate (μ)[12].

To examine the behavior of the system, a range of values for arrival rate (λ) and service rate (μ) are systematically selected. The values are chosen in such a way that the stability condition ($\lambda < \mu$) is satisfied for all scenarios. By varying λ while keeping μ constant, and vice versa, the study evaluates how changes in system input parameters influence idle time and overall system performance. This parametric analysis allows for a detailed investigation of system sensitivity and performance trends.

Besides theoretical analysis, simulation-based approach is also used to test the results obtained. Computational tools implemented to simulate include Python and Microsoft Excel that can be utilized to generate random arrival and service processes on the basis of exponential distributions. To achieve a control of consistency and reliability of results, several simulation runs are conducted with each combination of parameter values. Simulated values of idle time are next contrasted with theoretical evaluation to analyze the relevance and correctness of the model[13].

Moreover, the graphical analysis is the necessary element of the methodology to present the system behavior is drawn using line graphs.

The relationship between:

Idle time and arrival rate (λ)
Idle time and service rate (μ)

These graphical representations help in identifying trends, patterns, and correlations between variables, making the analysis more intuitive and interpretable[14].

A comparative analysis of theoretical and simulated results is also done to enhance the analysis. Any variations between the two are analyzed to know of possible causes of variability, e.g. randomness in simulation or approximation errors. Also, correlation is conducted to measure the magnitude and direction of relationship between arrival rate and idle time.

---

[9] Little, J. D. C. (1961). A fundamental relationship in queueing systems: The proof of L = λW. Operations Research, 9(3), 383–387.
[10] Cooper, R. B. (1981). Introduction to queueing theory (2nd ed.). New York, NY: North-Holland.
[11] Kleinrock, L. (1975). Queueing systems, volume 1: Theory. New York, NY: Wiley-Interscience.
[12] Hillier, F. S., & Lieberman, G. J. (2001). Introduction to operations research (7th ed.). New York, NY: McGraw-Hill.
[13] Medhi, J. (2003). Stochastic models in queueing theory (2nd ed.). San Diego, CA: Academic Press.
[14] Tijms, H. C. (2003). A first course in stochastic models. Chichester, UK: Wiley.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, April 2026 Impact Factor: 8.2

In general, this combined approach will be capable of providing a strong and stable examination of idle time in M/M/1 queueing system. The study has a powerful methodology in that by integrating mathematical modeling, simulation tools, and graphical explanations, the study builds a complete structure of understanding and optimizing the system performance.

VI. RESULTS AND DISCUSSION

The results obtained from both theoretical modeling and simulation provide a comprehensive understanding of idle time behavior in an M/M/1 queueing system. This section presents a detailed analysis of how variations in arrival rate (λ) and service rate (μ) influence system performance, particularly focusing on idle time, utilization, and system efficiency.

6.1 Numerical Results and Interpretation

To examine the relationship between arrival rate, service rate, and idle time, a set of numerical values was considered under the stability condition ($\lambda < \mu$). The results are summarized in Table 1.

Table 1: Idle Time Analysis for Different Arrival Rates

Arrival Rate (λ)	Service Rate (μ)	Utilization ($\rho = \lambda/\mu$)	Idle Time ($1 - \rho$)
2	5	0.40	0.60
3	5	0.60	0.40
4	5	0.80	0.20

The numerical results clearly demonstrate an inverse relationship between arrival rate (λ) and idle time. As λ increases from 2 to 4 while keeping μ constant at 5, the utilization factor (ρ) increases from 0.40 to 0.80. Consequently, idle time decreases from 0.60 to 0.20[15].

This behavior is consistent with theoretical expectations. When the arrival rate increases, more customers enter the system per unit time, thereby keeping the server occupied for a longer duration. As a result, the proportion of time during which the server remains idle reduces significantly.

From a practical standpoint, this indicates that systems with higher demand experience lower idle time but may face higher congestion if the arrival rate approaches the service rate.

6.2 Effect of Service Rate on Idle Time

To analyze the effect of service rate, μ was varied while keeping λ constant. The results are presented in Table 2.

Table 2: Idle Time Analysis for Different Service Rates

Arrival Rate (λ)	Service Rate (μ)	Utilization (ρ)	Idle Time
2	3	0.67	0.33
2	4	0.50	0.50
2	5	0.40	0.60
2	6	0.33	0.67

According to the results, the service rate (mu) and idle time have a direct correlation. The higher the μ the quicker the service is and the less time customers spend in the system. As a result, the server completes the service faster and spends more time when there are no other customers as soon as it does not receive any.

[15] Wolff, R. W. (1989). Stochastic modeling and the theory of queues. Englewood Cliffs, NJ: Prentice Hall.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, April 2026
Impact Factor: 8.2

This brings out a significant trade-off in system design. Although raising service rate enhances customer service and shortens the waiting time, it can result in insufficient resource utilization in case the demand is not high enough. As such, the best way to have the best performance of the system is by balancing the service capacity and the arrival rate.

6.3 Simulation vs Theoretical Comparison

To validate the theoretical findings, simulation was performed using computational tools. The comparison between theoretical and simulated idle time values is shown in Table 3.

Table 3: Simulation vs Theoretical Idle Time

Arrival Rate (λ)	Theoretical Idle Time	Simulated Idle Time
2	0.60	0.62
3	0.40	0.42
4	0.20	0.22

The theoretical results are close to the simulated values and there are a few deviations. These minor variations can be explained by the randomness of simulation processes as the arrival and services times are produced with the help of probabilistic distributions.

The similarity in theoretical and simulated outcomes proves that the M/M/1 model is valid and reliable in studying the behavior of idle time. It further shows that simulation is useful to create real-life situations where there exists variability and uncertainty.

6.4 Graphical Analysis

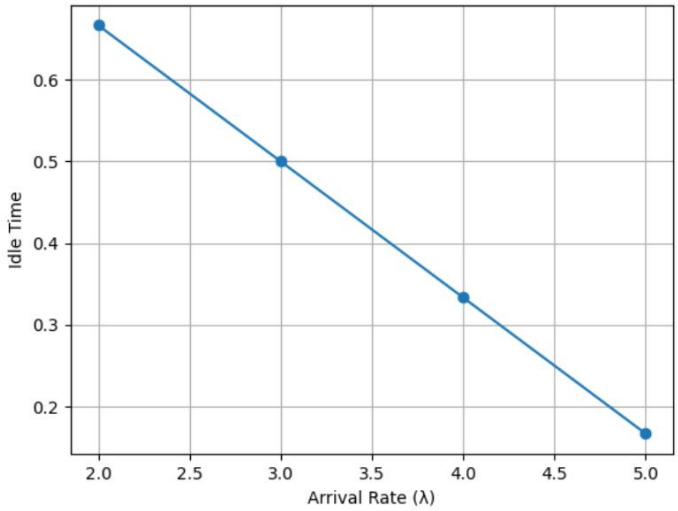

Figure 1: Idle Time vs Arrival Rate (λ)

The line graph of idle time and arrival rate has a downward slope which shows the negative relationship between the two. Idle time is reduced as λ increases. This ascertains the fact that the increased arrival rates translate to increased utilization of the servers and less idle time.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, April 2026
Impact Factor: 8.2

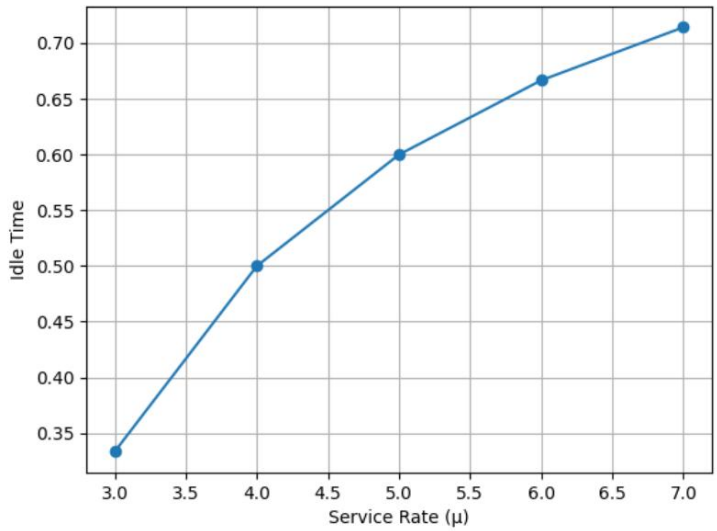

Figure 2: Idle Time vs Service Rate (μ)

The plot of the idle time versus the service rate is increasing and hence there is a positive relationship. The larger the μ, the more the idle time. This is because when the service is quick, the server gets to complete tasks much faster leaving it idle more often.

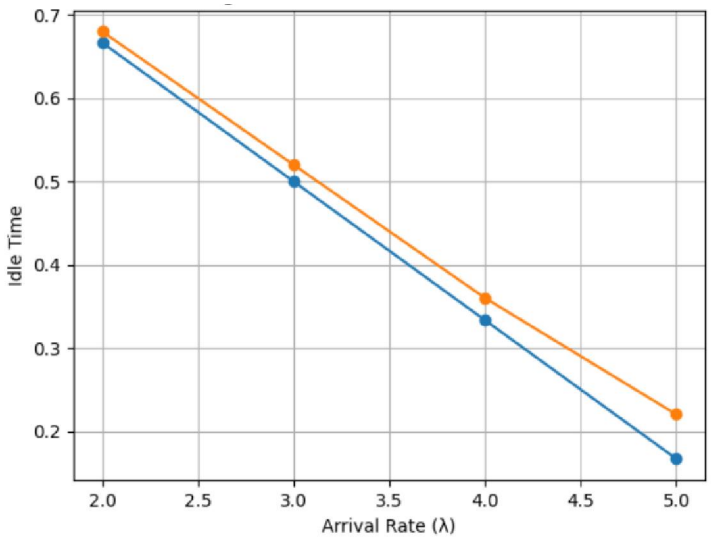

Figure 3: Simulation vs Theoretical Results

This graph illustrates that there are two curves that are very close in terms of theoretical and simulated idle time. The consistency of the similarity of the curves is a confirmation of the validity of theoretical formulations and shows that simulation outputs can be expected to follow the analytical expectations.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, April 2026
Impact Factor: 8.2

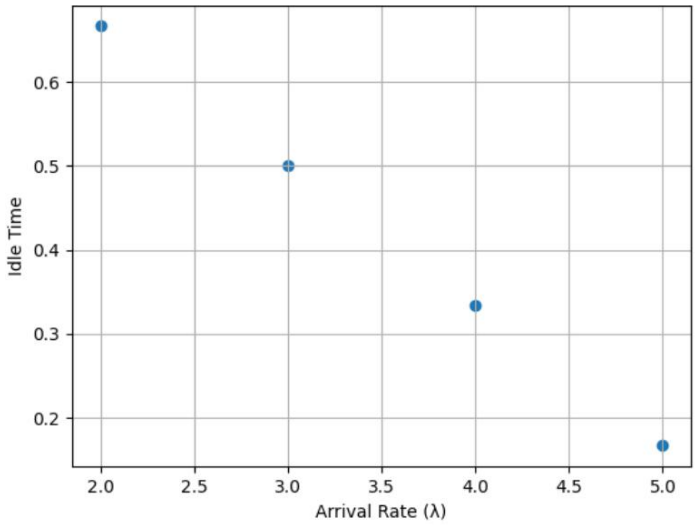

Figure 4: Correlation Analysis (λ vs Idle Time)

According to the scatter plot, there is a negative correlation between the arrival rate and idle time; its value is very high which means that the two are strongly correlated. This validates the fact that arrival rate is a prevailing factor affecting server usage and idleness.

6.5 Discussion of Key Findings

The findings of the research show the following significant ideas:

Negative Correlation between λ and Idle Time

The higher the arrival rate, the less idle time will be experienced by the server since the server will be in use serving customers.

Direct Relationship between mu and Idle Time

The service rates are higher resulting in more idle time since services are completed in a faster rate.

Stability and Usage of the System

When λ is very close to μ, the system is used up, and the time wasted is very minimal. This can however lead to overcrowding and longer waiting periods.

Efficiency versus Service Quality Trade-off.

Reduction of idle time enhances better utilization of resources but could have adverse effects on the time of waiting by customers. An increased idle time on the other hand guarantees faster service but underutilization.

Simultaneous verification with Simulation

The fact that the theoretical and simulated outcomes were very close proves that the M/M/1 model is an effective tool in the analysis of queueing systems.

VII. CONCLUSION

This paper will give a critical discussion of idle time in an M/M/1 queueing system. The findings show that the arrival rates, as well as service rates, have a considerable impact on idle time. The balance between μ and 2 should be maintained in order to achieve the best performance of the system.

The research proves that the higher the arrival rate the lesser the idle periods, whereas the higher the service rate the higher the idle periods. The fact that the theoretical and simulated outcomes are very close justifies the fact that the model is reliable.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, April 2026 Impact Factor: 8.2

The results can be used in practice to create effective service systems in telecommunications, healthcare, and manufacturing industries. The wasted time can be well managed, which results in better utilization of resources and cost efficiency.

REFERENCES

[1]. Madadi, M., Heydari, M. H., Maillart, L., Cassady, R., & Zhang, S. (2023). Erlang loss systems with shortest idle server first service discipline: Maintenance considerations. *IISE Transactions, 55*(10), 1008–1021.

[2]. Azhagappan, A., & Deepa, T. (2020). Variant impatient behavior of a Markovian queue with balking reserved idle time and working vacation. *RAIRO-Operations Research, 54*(3), 783–793.

[3]. Liu, Y., & Liu, B. (2021). Waiting time and idle time of uncertain queueing systems. *International Journal of General Systems, 50*(8), 871–890.

[4]. Ayyappan, G., Udayageetha, J., & Somasundaram, B. (2020). Analysis of non-preemptive priority retrial queueing system. *International Journal of Mathematics in Operational Research, 16*(4), 480–498.

[5]. Zhong, Y., Ward, A. R., & Puha, A. L. (2022). Asymptotically optimal idling in the GI/GI/N+GI queue. *Operations Research Letters, 50*(3), 362–369.

[6]. Erlang, A. K. (1909). *On the theory of probabilities in telephone traffic and congestion analysis*. Copenhagen: Nyt Tidsskrift for Matematik.

[7]. Kendall, D. G. (1953). Stochastic processes and their applications in queueing theory. *Journal of the Royal Statistical Society: Series B (Methodological), 15*(2), 151–185.

[8]. Gross, D., Shortle, J. F., Thompson, J. M., & Harris, C. M. (1998). *Fundamentals of queueing theory* (3rd ed.). New York, NY: Wiley.

[9]. Little, J. D. C. (1961). A fundamental relationship in queueing systems: The proof of L = λW. *Operations Research, 9*(3), 383–387.

[10]. Kleinrock, L. (1975). *Queueing systems, volume 1: Theory*. New York, NY: Wiley-Interscience.

[11]. Cooper, R. B. (1981). *Introduction to queueing theory* (2nd ed.). New York, NY: North-Holland.

[12]. Hillier, F. S., & Lieberman, G. J. (2001). *Introduction to operations research* (7th ed.). New York, NY: McGraw-Hill.

[13]. Medhi, J. (2003). *Stochastic models in queueing theory* (2nd ed.). San Diego, CA: Academic Press.

[14]. Tijms, H. C. (2003). *A first course in stochastic models*. Chichester, UK: Wiley.

[15]. Wolff, R. W. (1989). *Stochastic modeling and the theory of queues*. Englewood Cliffs, NJ: Prentice Hall.