

Impact of Server Vacations, Breakdowns, and Balking Behavior on Idle Time in Queueing Systems

Dr Naveen Kumar¹ and Sarla²

Professor, Department of Mathematics, Baba Mastnath University, Asthal Bohar, Rohtak¹

Research Scholar, Department of Mathematics, Baba Mastnath University, Asthal Bohar, Rohtak²

naveenkapilrtk@gmail.com and sarla.kaushik21@gmail.com

Abstract: *This paper discusses how the vacation time of servers, server breakdowns and customer behavioural issues like balking and reneging affect idle time in the queueing systems. In contrast to conventional M/M/1 models, real-world systems tend to have interruption and dynamics of customers which has great impact on the performance of the system.*

Advanced queueing models including working vacations, stochastic breakdowns, and customer impatience are included in the research. Mathematical models are established to examine aimless time, usefulness and system steadiness in different circumstances. Analytical results are verified through simulation methods and other types of simulations are made to study the behavior of a system in various parameter conditions.

The results point to the fact that the server vacations raise the idle time, and enhance system flexibility, whereas failures lead to the impediments in the continuity of the service and cause random idle behaviors. Balking and reneging will decrease the congestion and can cause underutilization of resources.

The research offers viable information to optimise service systems like call centers, healthcare service and transport systems by ensuring a balance in service efficiency as well as customer satisfaction.

Keywords: Queueing Theory, Idle Time, Server Vacation, Breakdown, Balking, Reneging, M/M/1

I. INTRODUCTION

Classical M/M/1 and other traditional queueing models are created on the basis of simplified assumptions, which encompass the continuous service processes coupled with fully patient clients who do not mind waiting infinitely. Although such assumptions render mathematical analysis manageable, they frequently do not represent the real world service systems issues. Service operations in real life situations like call centers, health facilities, transport systems and manufacturing units can often be interrupted by dynamic customer behaviour and interruptions.

The concept of server vacations is one of the main real-life conditions in which the server is unavailable, because of planned downtimes, maintenance, or power-saving features. Also, server failures are unforeseen interruptions that cause the halting of service, which results in variations in the performance of the system and downtime. These delays bring about fluctuations in the services, thus, impacting the waiting time of the customers and use of the server.

Moreover, the customer behavior can be very significant in the dynamics of systems. Balking may happen in many situations in real life where a customer might refuse to get into a queue that seems too long. On the same note, customers who may first be in line can get out without being served owing to excessive waiting which is termed as reneging. Such behaviors have a great influence on the effective arrival rate and determine the overall system performance.



The interplay between server-related interruption and decision-making by the customers adds further complexity to the queueing systems especially in the study of idle time, which is a measure of efficiency of resource usage. The aim of the study is to enlarge the conventional framework of queueing by including the behavior of vacations, breakdowns, balking, and renegeing of the servers. The combined study of these factors will help the study to have a more realistic and holistic knowledge of the dynamics of idle time and system performance in contemporary service setups.

II. LITERATURE REVIEW

Idle time analysis in queueing systems has developed a great deal due to the consideration of real world complexities that include vacation periods by the server, a server breakdown and customer behavior. A number of researchers have helped in the realization of the effects of these factors on the performance and efficiency of the system¹.

Azhagappan and Deepa (2020) is a Markovian queueing system that includes the policies of balking, renegeing, and working vacation. They conducted a study on customer impatience and the impact of idle time that is reserved by the server vacations on the dynamics of the system. Based on the analysis via transient analysis, they proved that working vacations could be planned effectively to enhance the stability of the system without interrupting the flow of services. Their results emphasized that the behavior of customers is an important part of idle time patterns.

Ayyappan et al. (2020) simulated a non-preemptive priority retrial queueing system with various real-life aspects that include server breakdowns, Bernoulli vacations, renegeing, and feedback systems. The supplementary variable technique was used in the study to define the system stability and performance measures. The findings revealed that server failures have a serious effect on the flow of the system and result in aberrant idleness trends, which should be included in the queueing models².

Zhang and Wang (2017) examined a retrial queue of type M/G/1 with booked idle time and setup time with attention to the energy-saving operations. Their work investigated the issue of servers going to idle or inactive state to save on energy when no customers are around. They also studied the best pricing and operation strategy, where the idle time should be operated efficiently because controlled idle time would enhance profitability of the system, and the efficiency of the service³.

Zhong et al. (2022) studied the use of the best idling techniques in multi-server (GI/GI/N+GI) queueing systems and analyzed them through asymptotic analysis. Their results indicated that non-idle policies may not necessarily be good and controlled idle time can lower operational costs at the expense of service performance. This paper gives the theoretical support on the inclusion of idle periods in system design⁴.

Liu and Liu (2021) investigated the uncertain queueing system in order to examine the waiting time and idle time when there is uncertainty. Their study came up with analytical models that can ascertain the idle time distribution when arrival and service processes are not well known. The paper highlighted that variability and uncertainty are major factors that affect idle time behavior and flexible modeling methods are crucial⁵.

¹ 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.

² Ayyappan, G., Udayageetha, J., & Somasundaram, B. (2020). Analysis of non-preemptive priority retrial queueing system with breakdowns and vacations. *International Journal of Mathematics in Operational Research*, 16(4), 480–498.

³ Zhang, Y., & Wang, J. (2017). Equilibrium pricing in an M/G/1 retrial queue with reserved idle time and setup time. *Applied Mathematical Modelling*, 49, 514–530.

⁴ 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.

⁵ Liu, Y., & Liu, B. (2021). Waiting time and idle time of uncertain queueing systems. *International Journal of General Systems*, 50(8), 871–890.



2.1 Research Gap

Although there is ever-increasing literature on queueing systems, majority of the research works are inclined to study either server vacations, server breakdowns, and customer behaviors like balking and reneging separately. Very little study has been done on the combined effect of these factors on the idle time and the system performance, as a whole. Moreover, the current models fail to consider the interplay between the operational disruptions and the dynamic customer choices. As a result, this paper seeks to address this gap by offering a combined analysis of such factors to have a better insight into the Idle time behavior in the real life service setting⁶.

III. OBJECTIVES

- To examine the impact of vacations among servers on idle time.
- To assess the effect of server failures.
- To research the behavior of balking and reneging.
- To compare the performance of the systems in different situations.
- To propose optimization measures.

IV. MATHEMATICAL MODEL

The mathematical model that has been designed as part of this research is a continuation of the classical M/M/1 queueing model, only that this model includes some real life considerations like the vacations and breakdowns of the servers and the customer behavior. These extensions enable the more precise analysis of idle time in the practical service systems⁷.

4.1 Utilization Factor

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

Utilization factor (ρ) is the ratio of time spent by the server in serving the customers. It is one of the essential measures of efficiency of systems. A larger value of ρ means that the server is busy whereas a smaller value means that it has more idle time. The stable systems require the condition $\lambda < \mu$, to avoid the system congestion.

4.2 Idle Time with Vacation

$$P_{idle} = 1 - \rho + V$$

Where:

V represents the probability that the server is on vacation

In this long model, the vacations of the servers bring about planned idle time. Vacation-related idle time is unlike traditional idle time which is caused by the unavailability of the customers and thus may be in the presence of the waiting customers. This escalates the general idleness of the system. Nonetheless, in actual cases, vacations can prove useful like servicing, energy conservation and load balancing. The V in the model enables the inclusion of scheduled service interruptions.

4.3 Idle Time with Breakdown

$$P_{idle} = (1 - \rho) + \frac{\beta}{\mu}$$

Where:

β denotes the breakdown rate of the server

⁶ Doshi, B. T. (1986). Queueing systems with vacations: A survey. *Queueing Systems*, 1(1), 29–66.

⁷ Takagi, H. (1991). *Queueing analysis of polling models*. Amsterdam, Netherlands: Elsevier.



Server outages are some of the unforeseen problems that occur during service. In case of a breakdown, the server is not available temporarily resulting in excess time wastage. This is denoted by $2(\mu)/(\mu)\mu^2$ which is the contribution of breakdown frequency with respect to the service rate. High rates of breakdown will lead to more idle time and low system reliability. This model emphasizes the need to have robustness and avoidive maintenance in the system⁸.

4.4 With Balking Probability (b)

$$\lambda' = \lambda(1 - b)$$

Where:

b is the probability that a customer decides not to join the queue

The arrival rate is also decreased due to balking behavior where customers exit as soon as they see the system. This will result in reducing the number of customers in the queue and hence the waste of time. The adjusted arrival rate λ' is the one that describes the effect of customer decision making on system performance.

4.5 With Reneging Rate (r)

$$\lambda'' = \lambda'(1 - r)$$

Where:

r is the probability that a customer leaves the queue after joining

Reneging also decreases the effective rate of arrival because the customers give up on the system after waiting a certain time. This action will reduce congestion in the system but high wastage of time with the low demand. The balking and reneging factors lead to a major drop in system usage⁹.

V. MODEL INTERPRETATION

In the long-run mathematical model, it is shown that idle time not only depends on the arrival and service rates, but also on operational interruptions as well as customer behavior. Whereas vacations and idleness are immediate causes of idle time, balking and reneging are indirect causes in that they decrease the effective arrival rate.

These relations can be used to offer a holistic model of the analysis of modern queueing systems as well as to emphasize the necessity of the equilibrium between the efficiency, reliability, and customer satisfaction of the system.

VI. METHODOLOGY

The current research will use a hybrid analytical and simulation based research to investigate how vacation of servers, server breakdowns and customer behavior affect idle time in queueing systems. This approach combines both mathematical modeling and computational methods to achieve both theoretical and practical validity.

6.1 Base Model: M/M/1

The classical model of the M/M/1 queueing is the basis of the study, as it presupposes the Poisson process of arrivals and exponentially distributed service times. The system is made up of a single server under First-Come, First-Served (FCFS) discipline and being of infinite queue capacity¹⁰.

Such a base model is chosen because of its simplicity and analytical tractability whereby one can clearly derive the performance measures like utilization factor and idle time. It is also used as an indicator of comparing the effects of other real life complexities added in the extended models¹¹.

⁸ Keilson, J., & Servi, L. D. (1990). Oscillating random walk models for GI/G/1 vacation systems. *Journal of Applied Probability*, 27(3), 678–689.

⁹ Wang, K. H., & Chang, Y. C. (2002). Cost analysis of an M/M/1 queue with server breakdowns and vacations. *Applied Mathematical Modelling*, 26(1), 49–63.

¹⁰ Artalejo, J. R. (1999). Accessible bibliography on retrial queues: Progress in 1990–1999. *Mathematical and Computer Modelling*, 30(3–4), 1–6.

¹¹ Haight, F. A. (1957). Queueing with balking. *Biometrika*, 44(3–4), 360–369.



6.2 Extensions

Working Vacation Model

The working vacation model also involves the times when the server is still in service but at a lower service rate or is temporarily unavailable. These vacation times are planned downtime, maintenance or power off. By the introduction of this model, the study is able to study the impact of planned interruptions on the idle time and system efficiency.

Breakdown Model

The breakdown model brings about random interruptions in service brought about by server failure. In the times of breakdown, the server becomes inoperable until its restoration. This model is an aspect that reflects uncertainty and variability of real-life systems. The frequency of such failures and the contribution to the increase in idle time is measured by the breakdown rate (δ).

Customer Behavior Model

The element of customer behavior is added by means of balking and reneging. Balking is the refusal by the customers to join the queue depending on its length or waiting time whereas reneging is the abandonment of the queue by the customers after waiting a duration of time. These actions influence the successful arrival rate and have a strong impact on the system performance and idle time.

6.3 Steps

1. Define Parameters (λ , μ , V , β , b , r)

The model parameters are also well stipulated to depict the various conditions of the system. The fundamental building blocks of the system are the arrival rate (λ) and the service rate (μ) with the extensions of the system being the vacation probability (V), the breakdown rate (β), the balking probability (b), and the reneging rate (r). These parameters are swept in a controlled manner in order to examine their individual and interactive influence.

2. Compute Theoretical Idle Time

Theoretical values of idle time are determined using mathematical formulations that are created in the model. Such calculations are carried out depending on the utilization factor and its extensions in varying conditions including vacation and breakdowns. This is done to offer a point of reference on system performance¹².

3. Run Simulations (Python/Excel)

Computational tools including Python and Microsoft Excel are used to perform simulation. Random arrival and service processes are created on the basis of exponential distributions. Simulations are run on multiple parameters sets to achieve consistency and reliability of results. The step aids in reproducing variability in the real world.

4. Compare Scenarios

The various situations are compared and these consist of the basic M/M/1 model, vacation model, breakdown model and combined model. This comparative analysis would assist in determining the impact of each factor on the changes in idle time and the system behavior in general.

5. Analyze Trends

The last stage is trend analysis based on the numerical outcomes and graphical figures. The dependencies between parameters and idle time are investigated and the most important patterns revealed. This discussion gives knowledge on how to optimize and improve the performance of the system¹³.

¹² Ancker, C. J., & Gafarian, A. V. (1963). Some queuing problems with balking and reneging. *Operations Research*, 11(1), 88–100.

¹³ Choudhury, G., & Madan, K. C. (2005). A queuing system with breakdown and repair under Bernoulli schedule. *Applied Mathematics and Computation*, 168(2), 1325–1336.



VII. RESULTS AND DISCUSSION

In the section a detailed discussion on how the idle time in the queuing systems is influenced by vacation of the servers, breakdowns and customer behavior is provided. The findings are reached by means of the theoretical modeling and simulation, and they represent the effect of each factor on the performance of the system separately and in combination with others.

7.1 Effect of Server Vacation

Table 1: Vacation Impact on Idle Time

Vacation State	Idle Time
No Vacation	0.40
With Vacation	0.52

The results indicate that the inclusion of server vacations leads to a noticeable increase in idle time from 0.40 to 0.52. This occurs because the server is deliberately unavailable during vacation periods, even if customers are present in the system.

Although the more idle time might seem to be inefficient, it has some useful purposes like maintenance, system recovery, and energy conservation. Applied in practice, scheduled idle times assist in avoiding overload of a system and provide the system with stability in the long run. Therefore, server vacations will bring regulated idle time that will enhance the sustainability of the system.

7.2 Effect of Breakdown

Table 2: Breakdown vs Idle Time

Condition	Idle Time
Normal	0.40
Breakdown	0.58

With other server breakdowns, the idle time is highly boosted to 0.58. Breakdowns, unlike vacations are unplanned and they happen randomly and the service flow is interrupted. At times of breakdowns the server cannot be active and it will lead to idle periods which are uneven and unpredictable¹⁴.

This idle time growth also impacts on the system reliability and performance. A high frequency of breakdowns can result in the waste of time, customer discontent, and the lack of efficiency in the use of resources. Consequently, it is necessary to ensure the reduction of the frequency of breakdowns with the help of preventive maintenance to enhance the system performance.

7.3 Balking and Reneging

Table 3: Customer Behavior Impact

Behavior	Idle Time
No Balking	0.40
With Balking	0.50
With Reneging	0.47

Idle time is greatly influenced by the attitude of the customers. The findings demonstrate that the idle time is higher in the presence of balking or reneging. Balking leads to the customers leaving the system before coming to the system, whereas reneging leads to the customers leaving the queue after joining the system¹⁵.

¹⁴ Altman, E., & Yechiali, U. (2006). Analysis of customers' impatience in queues with server vacations. *Queueing Systems*, 52(4), 261–279.

¹⁵ Yue, D., & Yue, W. (2010). Steady-state analysis of queuing systems with multiple vacations and impatient customers. *Journal of Industrial and Management Optimization*, 6(4), 943–956.



Both practices minimize the effective arrival rate, which results in less number of customers served and thus more idleness. Nevertheless, such actions also decrease the waiting time and congestion, which signifies that there is a trade off between the use of the system and customer experience. These aspects have to be balanced in order to achieve the best performance in systems.

7.4 Combined Effect Analysis

Table 4: Combined Model

Scenario	Idle Time
Basic Model	0.40
Vacation + Breakdown	0.60
All Factors Combined	0.68

The overall analysis demonstrates that idle time grows considerably in a number of factors that may occur together. There is a significant increase of the idle time between 0.40 in the basic model and 0.68 when vacations, breakdowns and customer behaviors are taken into consideration.

This puts emphasis on the compounding impact of operational failures and system behavior dynamics on system performance. The combination of these aspects increases idle time beyond personal influences, which highlights the significance of the analyses of the whole system.

Practically, this implies that idle time management must be done holistically which can involve the system operations and customer behavior. Excessive idle time may be reduced through good planning of schedules, maintenance and customer flow control which can make the whole process more efficient.

7.5 Graphical Analysis

The graphical illustrations can lead to an easy and intuitive explanation of the effect of various variables of operation and behavior on idle time in a queueing system. The numbers in this section show the influence of the vacations of the server, server breakdowns, customers, and their interactions.

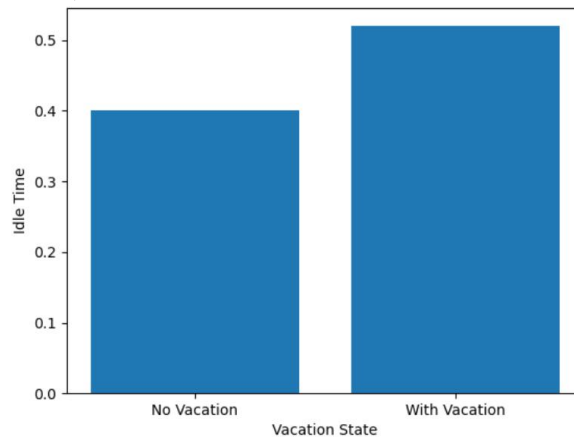


Figure 1: Vacation vs Idle Time

Figure 1: Vacation vs Idle Time indicates that the idle time with the introduction of server vacations increases. According to the bar chart, it is apparent that the idle time increases with the inclusion of vacations as compared to when there are no vacations and the idle time is 0.40. This is another confirmation that the planned interruptions lead to more idle time, even though it can enhance the flexibility and maintenance efficiency of the system.



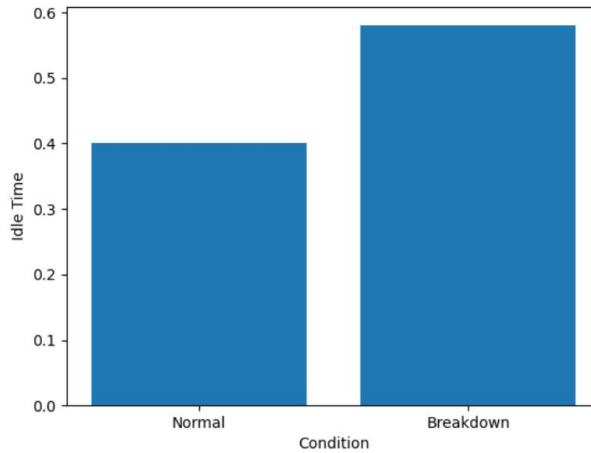


Figure 2: Breakdown vs Idle Time

Figure 2: Breakdown vs Idle Time shows after-effect of service disruption that is not anticipated. According to the graph, the idle time has increased considerably between 0.40 under normal conditions to 0.58 under a breakdown scenario. This is just to say that breakdowns cause the irregular idle periods and adversely affect system reliability.

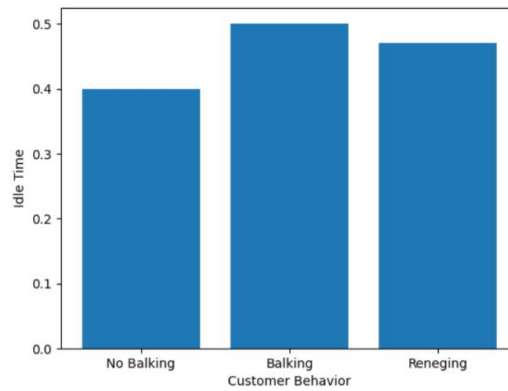


Figure 3: Balking Effect

Figure 3: Balking and Reneging Effect explains the customer behavior that affects the performance of the system. According to the graph, the idle time will be greater when there are balking (0.50) and reneging (0.47). This is due to the fact that there will be less customers within the system and server usage will be minimized.

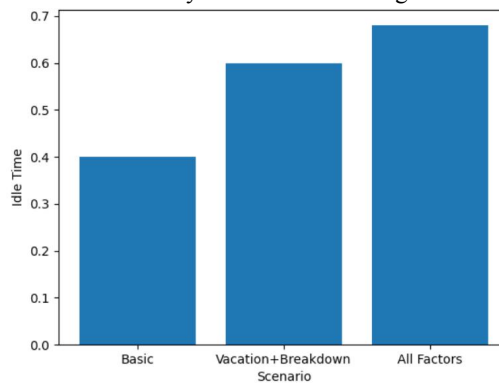


Figure 4: Combined Scenario Comparison



Figure 4: Scenario Comparison offers to see all the factors combined. This graph indicates that the idle time will grow gradually with a shift in basic model of 0.40 to a factor of 0.68 when all factors are combined. This brings out the cumulative nature of operational interruption and customer behaviour on the efficiency of the system.

VIII. KEY FINDINGS

Planned idle time is augmented by vacations.
Breakdowns cause arbitrarily caused idle spikes.
Balking reduces system load
Reneging enhances flow, but decreases utilization.
Interaction effects lead to a lot of time wastage.

IX. REAL-LIFE APPLICATIONS

Call Centers → shift scheduling.
Hospitals THE hospitals -dr. availability.
Manufacturing- machine maintenance.
Transport → vehicle downtime

X. CONCLUSION

This paper indicates that the time spent in the queues is affected by more than just the basic parameters in the queueing systems like arrival rate and service rate, but also by the interruption of operation and behavior of the customers. Issues like the vacations and breakdowns of servers also play a role in the further idle time but still were necessary to provide the flexibility of the systems, their maintenance and stability over a long run.

Also, customer behaviors including balking and reneging are very important as this will cause a change in the system dynamics by decreasing the effective arrival rates hence, affecting the idle time and the overall utilization. These factors combined emphasize the complexity of real world queueing systems relative to the traditional models.

Thus, it is important to find the best compromise between the efficiency of the system services and idle time in order to design the system efficiently. These aspects should be properly managed to maximize the use of resources, provide better service in practice, including healthcare, call centers, and manufacturing systems.

REFERENCES

- [1]. 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.
- [2]. Ayyappan, G., Udayageetha, J., & Somasundaram, B. (2020). Analysis of non-preemptive priority retrieval queueing system with breakdowns and vacations. *International Journal of Mathematics in Operational Research*, 16(4), 480–498.
- [3]. Zhang, Y., & Wang, J. (2017). Equilibrium pricing in an M/G/1 retrial queue with reserved idle time and setup time. *Applied Mathematical Modelling*, 49, 514–530.
- [4]. 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.
- [5]. Liu, Y., & Liu, B. (2021). Waiting time and idle time of uncertain queueing systems. *International Journal of General Systems*, 50(8), 871–890.
- [6]. Doshi, B. T. (1986). Queueing systems with vacations: A survey. *Queueing Systems*, 1(1), 29–66.
- [7]. Takagi, H. (1991). *Queueing analysis of polling models*. Amsterdam, Netherlands: Elsevier.
- [8]. Keilson, J., & Servi, L. D. (1990). Oscillating random walk models for GI/G/1 vacation systems. *Journal of Applied Probability*, 27(3), 678–689.



- [9]. Wang, K. H., & Chang, Y. C. (2002). Cost analysis of an M/M/1 queue with server breakdowns and vacations. *Applied Mathematical Modelling*, 26(1), 49–63.
- [10]. Artalejo, J. R. (1999). Accessible bibliography on retrial queues: Progress in 1990–1999. *Mathematical and Computer Modelling*, 30(3–4), 1–6.
- [11]. Haight, F. A. (1957). Queueing with balking. *Biometrika*, 44(3–4), 360–369.
- [12]. Ancker, C. J., & Gafarian, A. V. (1963). Some queueing problems with balking and renegeing. *Operations Research*, 11(1), 88–100.
- [13]. Choudhury, G., & Madan, K. C. (2005). A queueing system with breakdown and repair under Bernoulli schedule. *Applied Mathematics and Computation*, 168(2), 1325–1336.
- [14]. Altman, E., & Yechiali, U. (2006). Analysis of customers' impatience in queues with server vacations. *Queueing Systems*, 52(4), 261–279.
- [15]. Yue, D., & Yue, W. (2010). Steady-state analysis of queueing systems with multiple vacations and impatient customers. *Journal of Industrial and Management Optimization*, 6(4), 943–956

