

# Performance and Cost Analysis of Multi-Server Queues with Backup Servers under Server Failures

Satheeshkumar Ramadoss<sup>1</sup>, M.V.V.S. Nagendranath<sup>2</sup>, Srinivasan Rajavelu<sup>3</sup>, Parthiban Mohandas<sup>4</sup>,

Rakesh Kancharla<sup>5</sup>, E.V. Sandeep<sup>6</sup>

<sup>1</sup>Department of Mathematics, PSNA College of Engineering and Technology, Dindigul, Tamilnadu, India; [satheeshram74@gmail.com](mailto:satheeshram74@gmail.com)

<sup>2</sup>Department of Computer Science and Engineering, Sasi Institute of Technology and Engineering, Andhra Pradesh;

[nagendranath@sasi.ac.in](mailto:nagendranath@sasi.ac.in)

<sup>3</sup>Department of Electronics and Communication Technology, Sasi Institute of Technology and Engineering, Andhra Pradesh;

[srinivasan.researcher@gmail.com](mailto:srinivasan.researcher@gmail.com)

<sup>4</sup>Department of Computer Science and Engineering, Sasi Institute of Technology and Engineering, Andhra Pradesh; [parthimitit@gmail.com](mailto:parthimitit@gmail.com)

<sup>5</sup>Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh;

[rakesh51asstprof@gmail.com](mailto:rakesh51asstprof@gmail.com)

<sup>6</sup>Department of Computer Science and Engineering, Sasi Institute of Technology and Engineering, Andhra Pradesh; [evsandeep@sasi.ac.in](mailto:evsandeep@sasi.ac.in)

## ARTICLE INFO

## ABSTRACT

Received: 30 Dec 2024

Revised: 05 Feb 2025

Accepted: 25 Feb 2025

This study examines a multi-server tandem queueing system that is plagued by service disruptions from server failures and low service rates—problems that have a tendency to cause customer dissatisfaction, system delay, and abandonment. To counteract these impacts, we introduce an intelligent service continuity model with a backup server in case of primary server failures. Recognizing the cost-ineffectiveness of backup server activation for small queue sizes, the paper provides a strategic activation policy based on queue length, customer abandonment rate, and repair time. System dynamics are represented by a multi-dimensional continuous-time Markov process, which provides transient and steady-state behavior analysis. The model extends classical queueing systems by introducing Phase-Type (PH) distributions for the service and Markov Arrival Processes (MAPs) for accommodating correlated arrivals and sophisticated service behavior. There is not much in the literature that suggests extensive work has been done on tandem networks of such advanced stochastic modeling; Chakravarthy et al., Klimenok et al., and Dudin et al. have already developed basic methods, which are extended in the present work. Decision-theoretic techniques and cost-minimization functions are employed to determine optimal operating policies. Numerical experiments show the efficiency of employing strategic backup server operation in significantly reducing expected operational cost without compromising service quality, rendering the model a robust one for real-time mission-critical service systems.

**Keywords:** Tandem Queueing Networks, Standby Server, Customer Abandonment, Continuous-Time Markov Chain, Cost Optimization, Performance testing.

## INTRODUCTION

Tandem queueing networks, which are groups of several service stations, are extensively used in systems such as ambulance transport, manufacturing, telecommunications, and supply chains to improve service flow and reduce congestion. The systems are frequently plagued by service disruption due to failure or maintenance requirements of the servers, which greatly increases customer waiting time and abandonment risk. The addition of standby servers has proven to be an effective measure to handle this problem by providing continuity of service and system robustness. Madan et al. (2000) investigated an M/G/1 queue with optional secondary service in diverse operating conditions. Their work offered customer flow control and system stability insight through investigation of different service regimes and response mechanisms. Medhi et al. (2002) considered it important to study a priority M/G/1 queue with supplementary services. They used supplementary variable techniques to derive system performance measures like waiting time and queue length.

Wang et al. (2004) studied breakdown-and-feedback customer models in queues, such as analytical models of reliability analysis. They applied an approach of system availability analysis and downtime but with generalized conditions on service and arrivals. Choudhury et al. (2006) discussed in detail an M/G/1 queueing model for a two-phase service strategy. They were concerned with modeling optional service inclusion and transience of the queue. Ke et al. (2008) discussed how impatient customers influence an M/G/1 queue system when there is server failure and repair. They proposed models for incorporating customer behaviors like balking and reneging and repair time distribution. Muthukrishnan et al. (2010) developed a vacation policy model in M/G/1 queues with optional service. Their work enabled them to estimate system downtime and investigated policies for enhancing server availability. Singh et al. (2012) investigated an M[X]/G/1 vacation queue with state-dependent arrival rates and deterministic vacations. They employed supplementary variable techniques to derive queue length distribution and mean waiting time.

Chakravarthy et al. (2013) modeled a queueing framework with a common server advertising both obligatory and discretionary administrations. Their approach considered benefit prioritization and the affect on framework throughput. Jain et al. (2014) presented a bulk entry retrial line show with k-phase compulsory benefit and energetic get-away. This demonstrate was valuable for understanding complex entry designs and benefit interferences. Govindan et al. (2014) examined a retrial line with discretionary administrations and server disappointments beneath a Bernoulli get-away approach. They utilized supplementary variable investigation to assess client retry behavior and framework recuperation. Vijaya Laxmi et al. (2015) amplified the concept by modeling reneging behavior due to questionable administrations and coordination discretionary auxiliary benefit components. Abou-El-Ata et al. (1992) proposed an M/M/c line with client anxiety impacts. Their work highlighted client deserting amid expanded holds up and assessed framework capacity arranging. Xiong et al. (2009) analyzed a multi-server queueing framework for exchange handling with reneging interims and common benefit time. Their commitment lies in modeling settled restlessness limits in commonsense frameworks. Gomez-Corral et al. (2002) nursed a line display couple with an Outline entry handle and PH-distributed benefit times, with an infinite buffer at the main organize and none at the moment.

Van Houdt et al. (2005) estimated a discrete-time two-node couple line with PH benefit at hubs and bounded buffer confinements using matrix-analytic techniques. Lian et al. (2008) analyzed a tandem queue with PH service and infinite buffer capacities at both stations. Their work emphasized on capturing phase-type variability in service delivery. Baumann et al. (2013) considered tandem queues with MAP arrivals and PH service while modeling customer dropout after the first stage. Their work mimicked real-life multistage service environments. Kim et al. (2013) proposed a near-optimal solution to a PH service in stage two, a tandem queue-based call center with customer routing and overflow concerns. Singh et al. (2018) analyzed mandatory and voluntary service stages with stochastic models and complement variables to obtain steady-state probabilities upon server repair and maintenance.

Klimenok et al. (2018) derived an MMAP/PH/N standby server queueing model and solved the model in a multidimensional continuous-time Markov chain environment. Wang et al. (2018) analyzed congestion control with dynamic rate adaptation and modeled user choice behavior in heavy-load environments. Chakravarthy et al. (2021) employed level-dependent quasi-birth-and-death (LDQBD) processes to analyze group service models with PH and Weibull distributed general service times. Dudin et al. (2023) generalized this to multi-server systems where group service times depended on batch size and user patience. Their effort emphasized realism in modeling bulk service behavior. Dudin et al. (2024) further generalized this by adding group size constraints, which modeled service capacity constraints in actual systems. This study extends these works by incorporating a tandem queue with a standby server to assist the second stage of delivering mission-critical services. A value function is constructed to facilitate decision-making in the optimal sense in case of server failures to minimize the expected cost of the system and maximize customer retention by minimizing abandonment due to service disruption.

### MODEL DESCRIPTION

The Markov Choice Prepare may be a sound and common worldview of solutions for successive choice making beneath vulnerability, that's, optimization of choices inside an interminable skyline. The MDP relies on dynamic programming principles like state transition, optimality, and recursive value functions. Bellman first established the foundation for MDPs in 1957 with his principle of optimality, and subsequent work was derived later by Howard in

1960 using the policy iteration algorithm. MDPs are now very popular in optimization of operation research applied to a wide range of applications ranging from customer flow in service system, supply chain optimization, and dynamic inventory control. These models enable system designers to build policy-based, informed decisions aimed at optimizing resource and system performance. Of interest are research by authors such as Berman and Kim (1999), Bertsekas (2001), and Zayas-Cabán et al. (2016), which had set the broad applicability of MDPs in stochastic settings and demonstrated the utility of their use in the solution of operational issues in real life.

In order to simulate behavior that captures the dynamics of a service system, behavior may be simulated appropriately as a multidimensional continuous-time Markov chain. This applies to simulating stochasticity in state conversion and service processes, specifically in complex queuing systems. In order to simulate appropriately heterogeneous service times, multi-server configurations must be enforced with mandatory and discretionary phases of service. Exponential finds the broadest use in service time modeling because it has the benefits of memorylessness and analyzability by analysis if it is a result of queuing theory. PH distribution is, sometimes, used to model more realistic as well as heterogeneously observed service patterns. PH distribution has enough expressiveness to model any non-negative real-valued distribution and can potentially analyze those service processes with several phases or probabilistic phase transition. Its flexibility also renders it most suitable for heterogeneous service systems and more accurately portrays real service dynamics.

Client entry handle within the framework is Poisson and moment server repair time is exponential with cruel  $1/(\alpha)$ . One of the major challenges isn't permitting clients to renege amid the moment server repair time, and for this, a choice almost activating the backup server is required. For the most excellent utilization of the reinforcement server and decreasing the working taken a toll, the elective was not to switch it on on the simple rule that the littlest number of clients are online. This choice decreases the taken a toll of operations of the reinforcement server. In arrange to control the framework successfully and take rectify choices, the cost that happens in any state of the framework ought to be calculated and inspected. That's, considering the fetched of running the reinforcement server, for example, control utilization, operation and upkeep. It is of extraordinary significance to consider the assessed misfortunes within the frame of the disappointment and unwillingness of the client amid the period of repair. Based on cautious examination for all the states of the frameworks, the misfortunes with the two choices can be arranged appropriately. This estimation takes under consideration distinctive parameters just like the number of clients within the framework, repair time, benefit rates, client reneging likelihood and cost of running the reinforcement server.

The model is operating under Poisson passage uncertainty with  $\lambda$  rate. Mean time of the required server is very exponentially distributed in ruthless  $1/\mu_1$  and minute server mean time is PH distributed with parameters  $k$  and  $\mu_2$ . The minute server, being less in organizations, will have to get repaired periodically due to various unexpected reasons. Repair time of the minute server is exponential transport with parameter  $\alpha$ . Clients renege when the moment server is being repaired, and their reneging rate is  $qp$ , where  $p$  represents the likelihood that a reneged client is within the buffer of the moment server. Also, the remaining customers' normal holding up time at the moment server is  $1/\alpha q$ , with  $q=1-p$  being the likelihood that the client will stay holding up to be served at the moment server. In a offered to counter clog as well as reneging, the utilization of the reinforcement server can demonstrate to be profitable. Take note, however, that the reinforcement server is more expensive to function and less fast in benefit than the most server. The depiction to take after will uncover the way in which the state of the framework changes concurring to what activity has been taken. Figure (1) gives a couple lining framework with two servers and a limited holding up office for the customers. In expansion, a server is presented for reinforcement purposes to supply basic administrations at benefit rate  $\gamma$ , where  $\gamma$ .

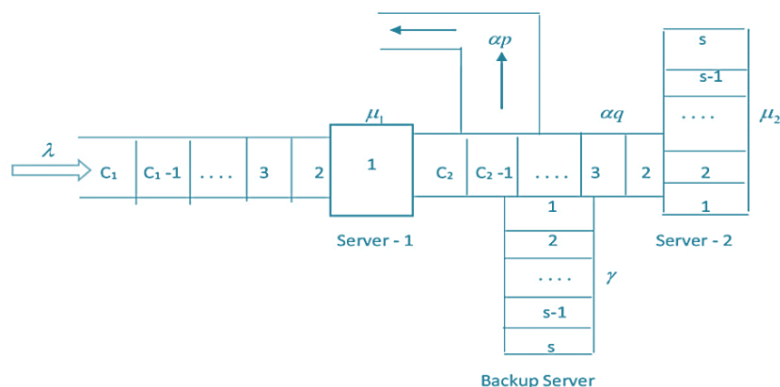


Figure 1: Tandem Queues with backup server in service facility system.

The goal is to have an optimum trade-off in such a manner that the total cost is kept minimum and customers are not hurt. Proper estimation and costing of the costs involved in each state should be possible. Decisions should be such that control over the system should be properly exercised so that the loss is minimum. The above situation can be probably best depicted through a queuing model that may be looked upon as an SMDP. In terms of different operating systems, system state information and state transition are necessary. State transitions under certain random and dynamic environments might be controlled through a series of operations.

### EXPERIMENTAL ANALYSIS

The complete gotten premise acknowledge that the strategy will definitely be wrapped up in a restricted number of steps. For the reason of comparing the execution of a course of action to an optimization premise, regard capacities  $V^\pi: S \rightarrow R$  are utilized.

Regard capacities separate the execution of an course of action  $\pi$  for each starting state  $i$  of the system. For each state of the system, the regard work gives the expected execution of the course of action  $\pi$  at anything point associated inside the state  $I$  and the MDP objective is to choose the perfect approach  $\pi^*$  due to that work  $v^*(\pi^*)$  (i) is reduced in this regard.

For all  $i \in I$  holds:  $(v^\pi(i)) \geq (v^*(\pi^*))(i)$

The taking after may be a depiction of checked down Markov choice issue theory and the put of the optimality condition and its course of action. The optimality condition joins a extraordinary course of action in  $V$ . The regard of the decreased Markov choice issue fulfills the optimality condition. The optimality condition depicts stationary perfect courses of action. There are perfect approaches underneath the fitting conditions for the states, exercises, costs, and move probabilities.

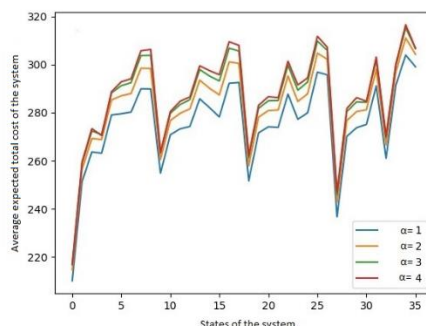


Figure 2: Average expected total cost for the different values of  $\alpha$

Because it is obvious from Fig. (2), with the rise within the repair rate of the moment server, the normal anticipated add up to taken a toll of the framework moreover rises, showing a positive relationship between the repair rate of the moment server and the overall taken a toll of the framework. Fig. (3) and Fig. (4) illustrate that the standard deviation

and cruel of the normal anticipated add up to taken a toll rise as there's a rise within the repair rate of the moment server. This demonstrates that, separated from normal fetched rises, anticipated costs are moving or spreading in an unexpected way beneath shifting circumstances. The discoveries above demonstrate the noteworthy impact of the repair time of the moment server on framework normal add up to anticipated fetched. With higher rate of repairing the moment server, there's higher taken a toll and higher inconstancy within the framework execution. This accounts for the need of legitimate administration and optimization of the repair handle of the moment server in arrange to decrease costs and make strides framework effectiveness.

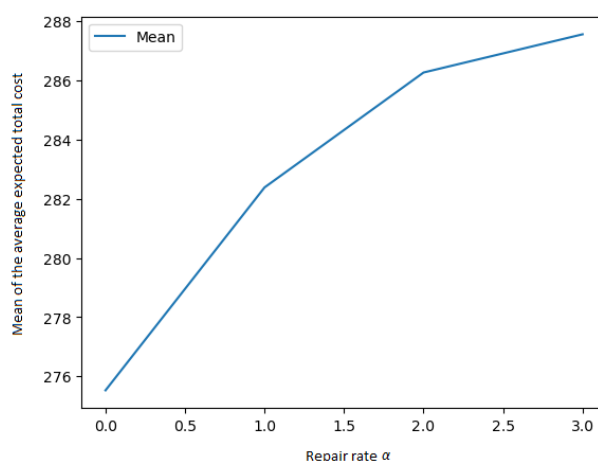


Figure 3: Mean of the average expected total cost for different values of  $\alpha$ .

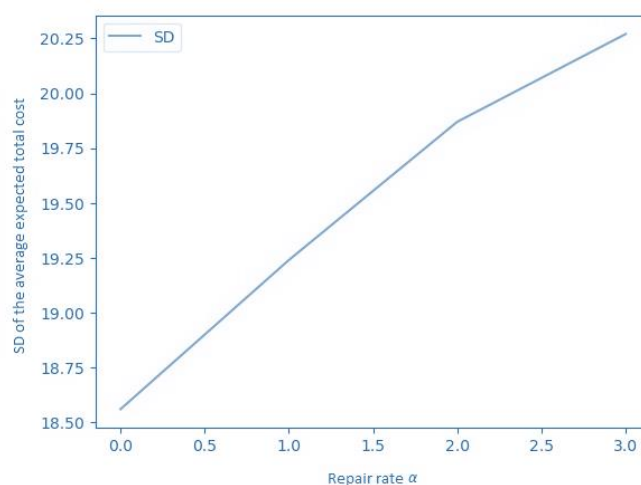


Figure 4: SD of the average expected total cost for different values of  $\alpha$ .

Ideal normal achieved for the diversified values of  $\alpha$  and  $p$ : Ideal normal achieved for the diversified values of  $\alpha$  provided in Table (5) is calculated using esteem cycle, whereby the ideal normal accumulated a cost of the framework.

$\alpha$	$p$	Optimal Avg.cost	No. of iterations
1	0.2	13.20	48
	0.4	13.48	48
	0.5	13.61	48
	0.6	13.73	49
	0.8	13.65	51
2	0.2	13.29	47
	0.4	13.85	48
	0.5	14.05	49
	0.6	14.22	49

3	0.8	13.69	47
	0.2	13.36	48
	0.4	14.16	49
	0.5	14.41	50
	0.6	14.22	49
	0.8	13.51	50
4	0.2	13.41	49
	0.4	14.41	49
	0.5	14.47	50
	0.6	14.15	51
	0.8	13.03	52

Table 5: Maximum average long-run cost for all possible values of  $\alpha$  and  $p$ .

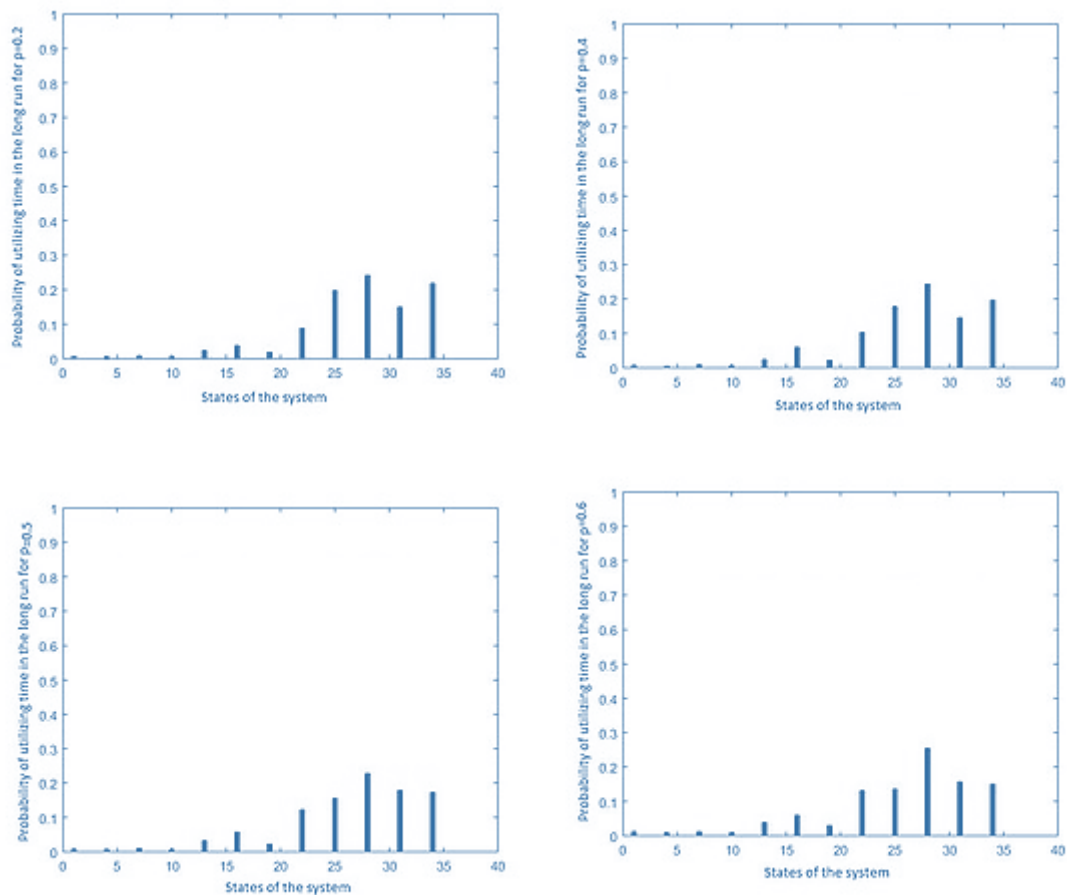
Unfaltering state dispersion in each state of the framework: Framework examination for different probabilities of reneging  $p$  and assurance of the stationary likelihood disseminations. Utilize of time in each state of the framework can be diverse based on the given parameter. For, being fixed in the long run for the specific values of  $p$  is provided by Table (6) and framework states emerged in Fig. (5) are realizable in the long run.

States	$p=0.2$	$p=0.4$	$p=0.5$	$p=0.6$	$p=0.8$
$I_1$	0.0069	0.0073	0.0081	0.0126	0.0269
$I_2$	0.0000	0.0000	0.0000	0.0000	0.0002
$I_3$	0.0000	0.0000	0.0000	0.0000	0.0000
$I_4$	0.0048	0.0052	0.0073	0.0094	0.0150
$I_5$	0.0000	0.0000	0.0000	0.0000	0.0006
$I_6$	0.0000	0.0000	0.0000	0.0000	0.0000
$I_7$	0.0083	0.0087	0.0097	0.0120	0.0163
$I_8$	0.0000	0.0000	0.0000	0.0000	0.0010
$I_9$	0.0000	0.0000	0.0000	0.0000	0.0000
$I_{10}$	0.0055	0.0058	0.0065	0.0097	0.0173
$I_{11}$	0.0000	0.0000	0.0000	0.0000	0.0009
$I_{12}$	0.0000	0.0000	0.0000	0.0000	0.0000
$I_{13}$	0.0228	0.0240	0.0317	0.0390	0.0475
$I_{14}$	0.0000	0.0000	0.0000	0.0000	0.0038
$I_{15}$	0.0000	0.0000	0.0000	0.0000	0.0000
$I_{16}$	0.0380	0.0597	0.0574	0.0594	0.0473
$I_{17}$	0.0000	0.0000	0.0000	0.0000	0.0111
$I_{18}$	0.0000	0.0000	0.0000	0.0000	0.0000
$I_{19}$	0.0191	0.0220	0.0225	0.0295	0.0421
$I_{20}$	0.0000	0.0000	0.0000	0.0000	0.0011
$I_{21}$	0.0000	0.0000	0.0000	0.0000	0.0000
$I_{22}$	0.0883	0.1033	0.1226	0.1308	0.1311
$I_{23}$	0.0000	0.0000	0.0000	0.0000	0.0037
$I_{24}$	0.0000	0.0000	0.0000	0.0000	0.0000
$I_{25}$	0.1967	0.1777	0.1555	0.1356	0.0943
$I_{26}$	0.0000	0.0000	0.0000	0.0000	0.0060



$I_{27}$	0.0000	0.0000	0.0000	0.0000	0.0000
$I_{28}$	0.2417	0.2432	0.2280	0.2542	0.3085
$I_{29}$	0.0000	0.0000	0.0000	0.0000	0.0012
$I_{30}$	0.0000	0.0000	0.0000	0.0000	0.0000
$I_{31}$	0.1494	0.1457	0.1780	0.1571	0.1330
$I_{32}$	0.0000	0.0000	0.0000	0.0000	0.0026
$I_{33}$	0.0000	0.0000	0.0000	0.0000	0.0000
$I_{34}$	0.2185	0.1974	0.1728	0.1507	0.0857
$I_{35}$	0.0000	0.0000	0.0000	0.0000	0.0028
$I_{36}$	0.0000	0.0000	0.0000	0.0000	0.0000
$\sum I_i$	1.0000	1.0000	1.0000	1.0000	1.0000

Table 6: The long run steady state distribution of the system states for the various values of p, where=4.



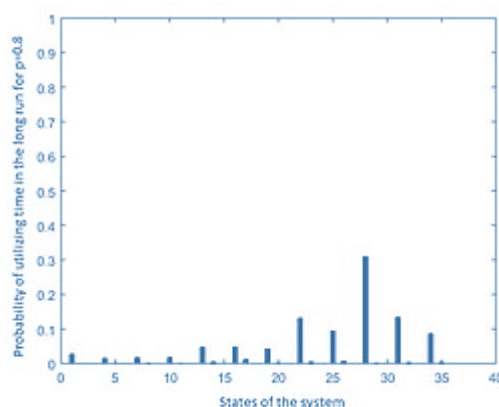


Figure 5: Time rate, the state of the framework is used in a long run for various values of  $p$  when  $\alpha=4$ .

Administrative suggestions: In a multi-server lining framework having a standby server, utilization of the standby server in times of require is noteworthy in giving best benefit. This is often at the same time required to be adjusted with minimizing the by and large taken a toll. For the reason, the effect of putting the standby server is based on the number of clients that exit the holding up room when the moment server is in require of repair. The moment server repair time and client renege likelihood both contribute to the long-run framework taken a toll. Both of them are appeared in Table (3) and Table (4), whereas their graphical representations can be watched from Fig. (2), Fig. (3) and Fig. (4). The ideal normal taken a toll of the entire framework is computed in Table (5) for various values of  $\alpha$  and  $p$ . On the off chance that we see at this table, we are able naturally watch that the ideal normal fetched rises within the interim of for  $\alpha = 1$  and  $\alpha = 2$ . On the other hand, the ideal normal costs drop in the interim of for  $\alpha = 3$  and  $\alpha = 4$ . In expansion: Table (6) gives the likelihood of time utilize in each state of the framework, which is outlined graphically in Fig. (5).

## CONCLUSION

The current ponder takes a multi-server couple lining framework with limited holding up space to be served in a Markov choice demonstrate. The demonstrate obliges a number of taken a toll parameter, counting holding, preparing, holding up, renege, and benefit costs. The anticipated add up to taken a toll and ideal normal taken a toll are calculated utilizing the esteem cycle calculation. The investigation found that the ideal normal taken a toll and add up to anticipated taken a toll both rise with repair time and client renege likelihood. Be that as it may, employing a reinforcement server decreases these costs. The discoveries give bits of knowledge into the system's behavior and execution. Future investigates seem investigate changing over the show into a parallel line demonstrate, considering reinforcement servers amid repair or get-away periods, and renege and balking clients. The In part Watched Markov Choice Handle (POMDP) system can be utilized to analyze framework execution measures.

## REFERENCES

- [1] M. Abou-El-Ata and S. Ibrahim, "Two-servers heterogeneous overflow queues," J. King Abdulaziz Univ. Sci., vol. 4, pp. 159–168, 1992.
- [2] O. Alagoz and M. Ayvaci, "Uniformization in Markov Decision Processes," Wiley Online Library, 2011.
- [3] H. Baumann and W. Sandmann, "Computing stationary expectations in level-dependent QBD processes," J. Appl. Probab., vol. 50, pp. 151–165, 2013.
- [4] R. E. Bellman, Dynamic Programming, Princeton University Press, 1957.
- [5] O. Berman and E. Kim, "Stochastic models for inventory management at service facilities," Commun. Stat. Stochastic Models, vol. 15, pp. 695–718, 1999.
- [6] D. P. Bertsekas, Dynamic Programming and Stochastic Control, Athena Scientific, 2001.
- [7] S. R. Chakravarthy, "Analysis of MAP/PH1, PH2/1 queue with vacations and optional secondary services," Appl. Math. Modell., vol. 37, pp. 8886–8902, 2013.



- [8] G. Choudhury and M. Paul, "A batch arrival queue with a second optional service channel under n-policy," *Stochastic Anal. Appl.*, vol. 24, pp. 1–21, 2006.
- [9] S. R. Chakravarthy, G. Shruti, and A. Rummyantsev, "Analysis of a queueing model with batch Markovian arrival process and general distribution for group clearance," *Methodol. Comput. Appl. Probab.*, vol. 23, pp. 1551–1579, 2021.
- [10] S. Dudin and O. Dudina, "Analysis of a multi-server queue with group service and service time dependent on the size of a group as a model of a delivery system," *Mathematics*, vol. 11, no. 22, p. 4587, 2023. [Online]. Available: <https://doi.org/10.3390/math11224587>.
- [11] S. A. Dudin, O. S. Dudina, and A. N. Dudin, "Analysis of tandem queue with multi-server stages and group service at the second stage," *Axioms*, vol. 13, no. 4, p. 214, 2024. [Online]. Available: <https://doi.org/10.3390/axioms13040214>.
- [12] A. Gomez-Corral, "A tandem queue with blocking and Markovian arrival process," *Queueing Syst.*, vol. 41, pp. 343–370, 2002.
- [13] A. Govindan and S. Shyamala, "M[X]/G/1 retrial queueing system with second optional service, random breakdown, set up time and Bernoulli vacation," *Int. J. Open Probl. Comput. Sci. Math.*, vol. 7, pp. 23–39, 2014.
- [14] Krishnamoorthy, M.P., 2024. Big Data Analytics In Fintech: A Review Of Credit Risk Assessment And Fraud Detection. *Educational Administration: Theory and Practice*, 30 (5), pp.3676-3684.
- [15] D. Heyman and M. Sobel, *Stochastic Models*, Elsevier Science Publications, 1990.
- [16] R. A. Howard, *Dynamic Programming and Markov Processes*, John Wiley and Sons, Inc., 1960.
- [17] M. Jain and A. Bhagat, "Unreliable bulk retrial queues with delayed repairs and modified vacation policy," *J. Ind. Eng. Int.*, vol. 10, pp. 1–19, 2014.
- [18] J.-C. Ke, "Two thresholds of a batch arrival queueing system under modified t-vacation policy with startup and closedown," *Math. Meth. Appl. Sci.*, vol. 31, pp. 229–247, 2008.
- [19] C. Kim, A. Dudin, S. Dudin, and O. Dudina, "Tandem queueing system with impatient customers as a model of call center with interactive voice response," *Perform. Eval.*, vol. 70, pp. 440–453, 2013.
- [20] V. Klimenok, A. Dudin, and K. Samouylov, "Analysis of the BMAP/PH/N queueing system with backup servers," *Appl. Math. Modelling*, vol. 57, pp. 64–84, 2018.
- [21] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*, Society for Industrial and Applied Mathematics, 1999.
- [22] Deepak, G., Parthiban, M., Nath, S.S., Alfurhood, B.S., Mouleswararao, B. and Kishore, V.R., 2024. Ai-enhanced thermal modeling for integrated process-product-system optimization in zero-defect manufacturing chains. *Thermal Science and Engineering Progress*, 55, p.102945.
- [23] Z. Lian and L. Liu, "A tandem network with MAP inputs," *Oper. Res. Lett.*, vol. 36, pp. 189–195, 2008.
- [24] D. M. Lucantoni, K. Meier-Hellstern, and M. F. Neuts, "A single server queue with server vacations and a class of non-renewal arrival processes," *Adv. Appl. Probab.*, vol. 22, pp. 676–705, 1990.
- [25] K. C. Madan, "An M/G/1 queue with second optional service," *Queueing Syst.*, vol. 34, pp. 37–46, 2000.
- [26] J. Medhi, "A single server Poisson input queue with a second optional channel," *Queueing Syst.*, vol. 42, pp. 239–242, 2002.
- [27] S. K. M. Muthukrishnan and R. Arumuganathan, "An M[X]/G/1 retrial queue with two-phase service subject to active server breakdowns and two types of repairs," *Int. J. Oper. Res.*, vol. 8, pp. 261–291, 2010.
- [28] M. Neuts, "A versatile Markovian point process," *J. Appl. Probab.*, vol. 17, pp. 764–799, 1979.
- [29] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models*, Johns Hopkins University Press, 1981.
- [30] C. A. O’Cinneide, "Characterization of phase-type distributions," *Commun. Stat. Stochastic Models*, vol. 6, pp. 1–57, 1990.
- [31] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley and Sons, Inc., 1994.
- [32] Rao, B. S., Parthiban, M., Sugumaran, D., Beevi, S. Z., Suganya, C., & Manikandan, G. (2024). Integrating Transfer Learning with Multimodal Mri Fusion for Enhanced Detection of Multiple Sclerosis Lesions. In *Journal Of Environmental Protection and Ecology* (Vol. 25, Issue 8, Pp. 2790–2799).

- [33] R. Sabbadin, D. Spring, and C.-E. Rabier, "Dynamic reserve site selection under contagion risk of deforestation," *Ecol. Modelling*, vol. 201, pp. 75–81, 2007.
- [34] C. J. Singh, M. Jain, and S. Kaur, "Performance analysis of bulk arrival queue with balking, optional service, delayed repair and multi-phase repair," *Ain Shams Eng. J.*, vol. 9, no. 4, pp. 2067–2077, 2018.
- [35] C. J. Singh, M. Jain, and B. Kumar, "Analysis of M/G/1 queueing model with state dependent arrival and vacation," *J. Ind. Eng. Int.*, vol. 8, pp. 1–8, 2012.
- [36] H. Tijms, *Stochastic Models: An Algorithmic Approach*, Wiley, New York, 1994.
- [37] B. Van Houdt and A. S. Alfa, "Response time in a tandem queue with blocking, Markovian arrivals and phase-type services," *Oper. Res. Lett.*, vol. 33, pp. 373–381, 2005.
- [38] P. Vijaya Laxmi and K. Jyothsna, "Balking and reneging multiple working vacations queue with heterogeneous servers," *J. Math. Modelling Algorithms Oper. Res.*, vol. 14, pp. 267–285, 2015.
- [39] J. Wang, "An M/G/1 queue with second optional service and server breakdowns," *Comput. Math. Appl.*, vol. 47, pp. 1713–1723, 2004.
- [40] Gokulkannan, K., Parthiban, M., Jayanthi, S. and Kumar, M., 2024. Cost effective cloud-based data storage scheme with enhanced privacy preserving principles. *The Scientific Temper*, 15(02), pp.2104-2115.
- [41] Q. Wang and B. Zhang, "Analysis of a busy period queueing system with balking, reneging and motivating," *Appl. Math. Modelling*, vol. 64, pp. 480–488, 2018.
- [42] W. Xiong and T. Altiok, "An approximation for multi-server queues with deterministic reneging times," *Ann. Oper. Res.*, vol. 172, pp. 143–151, 2009.
- [43] G. Zayas-Cabán, J. Xie, L. V. Green, and M. E. Lewis, "Dynamic control of a tandem system with abandonments," *Queueing Syst.*, vol. 84, pp. 279–293, 2016.