

Study on Bernoulli scheduling, customers priority shift in busy state, intolerance of customers in working vacation, complete vacation in a $M/M/1/\infty$ model

Rachna Rathore ¹, R K Shrivastava ²

Abstract: This study examines Bernoulli scheduling approach, customers(clients) priority shift during server's busy state and explores customers intolerance during server's working vacation (WV) state with complete vacation in a $M/M/\infty$ queueing model. During the busy state, customers might join the queue with a probability d or leave with a probability $s = 1 - d$. Some customers might enter the system but do not join the queue during the busy state are assumed to result from changing priorities or other pressing matters, rather than due to long wait times or frustration. The server enters a working vacation in free time with an exponentially distributed rate θ , providing service at a slower pace. During this phase, customers may become impatient if faced with extended waits, abandoning the queue with an exponentially distributed parameter α if their patience expires. Upon WV completion, the server either resumes a busy state with probability p if any customer present, or join complete vacation with probability m and rate β . If new clients arrive during complete vacation, the server returns to busy state at a rate ψ otherwise availing complete vacation. This paper employs the PGF method to derive equilibrium-state probabilities and assess system measures, including average queue lengths during busy state, WV states and complete vacation, mean sojourn time, and abandonment rate. Additionally, the influence of specific parameters on system metrics is illustrated graphically.

Keywords: Busy period; Bernoulli schedule; Priority shift; Working vacation; Impatience of customers; Rate of abandonment of customers

2020 Mathematics Subject Classification: 60K25

Receive: 10 July 2024, **Accepted:** 12 November 2024

1 Introduction

In recent years, researchers have extensively explored the application of various queuing models across numerous fields. Queuing systems with server vacations are increasingly relevant in scenarios where server may occasionally be unavailable due to fatigue or other reasons. By integrating server vacation periods, queuing models provide a framework for optimizing processes, thereby balancing resource utilization and service efficiency. However, in some cases, the server continues to operate at a reduced pace during vacation

¹Corresponding author: Research Scholar, Shrimant Madhav Rao Scindia Govt. Model Science College, Jiwaji University, Gwalior, M.P. India, Email: rachnarathore1109@gmail.com

²Principal, Dr. Bhagwat Sahay Gov. College, Gwalior, M.P., India

rather than halting entirely. The concept of a working vacation allows the server to maintain service at a slower rate, often used when demand is lower to improve resource use. Servi et al. [23] introduced the idea of single and many working vacations within a one-server Markovian queue. Following this, Liu et al. [16] and Tian et al. [26] derived stochastic decompositions for single-server queues with one or multi working vacations respectively. Lin et al. [15] extended Tian et al.'s work to multi-server systems. Selvaraju and Goswami [22] examined client impatience in one and multiple WV settings.

In daily life, situations often arise where customers become impatient while waiting in line. This impatience can lead customers to avoid joining the queue altogether (balking) or to leave after joining but before receiving service (reneging). Haight [9] first introduced balking in one server queues, later expanding on reneging in Markovian systems [10]. Ancker et al. [3, 4] examined the impact of both behaviors in finite-length single-server queues, while Altman and Yechiali [2] studied customer intolerance during periods of server unavailability. Yechiali [28] further analyzed impatience and system failures in downtime situations, and Perel and Yechiali [18] investigated client impatience in variable Markovian environments. Yue et al. [29] studied impatience during working vacations in a single-server queue, while Goswami [7] analyzed impatience and vacation interruptions within a Bernoulli-scheduled Markovian queue with multi working vacations. Laxmi et al. [13] gave a detailed analysis of a one-server queue featuring WV with customer balking. Yang et al. [27] examined client retention in a finite Markovian queue with server breakdowns(malfunction) and repairs. Swathi et al. [25] researched reneging behavior across different server states in a one server Markovian system with one and many vacations. Rathore [19, 20] notably investigated queueing models within the healthcare sector and utilizing stochastic modelling to manage crowding and congestion respectively. Devi et al. [6] evaluated a one-server Markovian model incorporating server deterioration and customer intolerance. Ahmed et al. [1] explored retention and state-dependent reneging within an $M/M/1/\infty$ system.

Shrivastava & Rathore [24] determined several system metrics and equilibrium state probabilities in a one server Markovian queue with reneging of customers in differentiated working vacation.

The Bernoulli scheduling policy in queuing theory describes a server's decisionmaking process on whether to take a vacation or serve a customer, guided by a Bernoulli process. Keilson et al. [11] were pioneers in studying Bernoulli-scheduled service, which was later expanded by Kella [12] to include a generalized model where a server can take successive vacations based on a probability-based scheme. Choudhury and Paul [5] applied Bernoulli scheduling to a two-phase system with multiple vacations, while Li et al. [14] examined reliability in models with Bernoulli vacations. Zhang et al. [30] discussed controlled Bernoulli scheduling with vacation interruptions in an $M/M/1$ queue, and Manoharan and Jeeva [17] studied customer impatience in a Bernoulli-scheduled, working vacation model with interruptions and setup times by using PGF method. Gupta and Kumar [8] investigated a Bernoullischeduled, $M/M/1$ queue with interruption in working vacation, focusing on customer balking behavior. This paper extend work of Rathore and Shrivastava [21].

2 Model Description

This study considers an $M/M/1$ queueing model characterized a Bernoulli scheduling policy, customers varying or shifting priorities in busy state, customers intolerance during the working vacation (WV) state with complete vacation. Customers arrival following a Poisson distribution with an arrival rate λ , and service times adhere to an exponential distribution with parameter μ . During the busy state, clients may enter or exit the queue with probabilities d and $s = 1 - d$ respectively. It is expected that those clients who approach the system but do not join the queue during busy state do so because their priorities have changed or they have other urgent items to attend to, rather than because they are frustrated or have to wait a long time. When the system is empty, the server transitions to a working vacation, characterized by an exponentially distributed rate θ . In this state, the server operates at a reduced service rate η . During the

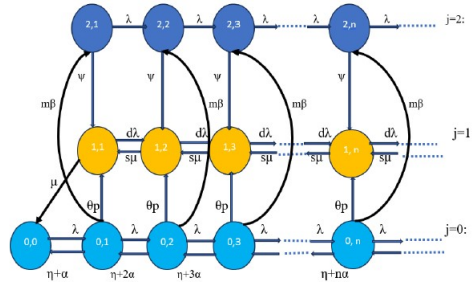


Figure 1: Transition rate diagram of the Model.

working vacation, customers are considered impatient, with an exponentially distributed intolerance rate α owing to the slower pace of service. If a customer’s service does not commence before their impatience timer expires, they will abandon the system permanently. Customers who have already begun receiving service, however, will remain in the system until completion. Only those waiting in the queue during the working vacation phase are susceptible to becoming impatient and leaving. Upon the completion of the working vacation, the server resumes a busy state with probability p if at least one customer is present in the system. If there are no customer waiting, the server joins complete vacation with probability m and rate β . If a new customer arrives while the server is in complete vacation, the server will revert to the busy state with an exponential rate ψ . All clients are served on FCFS basis. A transition diagram of the server states is shown in Figure 1.

3 Mathematical Formulation

Let $N(t)$ = Number of customers in the system at time t ,

$J(t)$ = State of server at time t with

$j = 0$ denotes server is in working vacation state

$j = 1$ denotes server is in busy state.

$j = 2$ denotes server is in complete vacation.

Consider a CTMC process $(J(t), N(t))$ with transition rate diagram as shown in Figure 1.

Let $P_{jn} = P\{J(t) = j, N(t) = n; j = 0, 1, 2; n = 0, 1, 2, \dots\}$ denotes the equilibrium state probabilities of the system.

Balance equations for $j = 0$:

$$\lambda P_{00} = (\alpha + \eta)P_{01} + \mu P_{11}, \quad n = 0 \quad (3.1)$$

$$(\lambda + n\alpha + \eta + \theta p + m\beta)P_{0n} = \lambda P_{0,n-1} + [\eta + (n + 1)\alpha]P_{0,n+1}, \quad n \geq 1 \quad (3.2)$$

Balance equations for $j = 1$:

$$(d\lambda + \mu)P_{11} = \psi P_{21} + s\mu P_{12} + \theta p P_{01}, \quad n = 1 \quad (3.3)$$

$$(d\lambda + s\mu)P_{1n} = \psi P_{2n} + s\mu P_{1,n+1} + \theta p P_{0n} + d\lambda P_{1,n-1}, \quad n \geq 2 \quad (3.4)$$

Where $s = 1 - d$.

Balance equations for $j = 2$:

$$P_{21}(\lambda + \psi) = m\beta P_{01}, \quad n = 1 \quad (3.5)$$

$$P_{2n}(\lambda + \psi) = m\beta P_{0n} + \lambda P_{2n-1}, \quad n \geq 2 \quad (3.6)$$

Now we use PGF technique to derive various system metrics and equilibrium state probabilities,

$$\text{Let } I_0(z) = \sum_{n=0}^{\infty} P_{0n} z^n, \quad I'_0(z) = \sum_{n=0}^{\infty} n P_{0n} z^{n-1} \quad (3.7)$$

$$I_1(z) = \sum_{n=1}^{\infty} P_{1n} z^n, \quad I'_1(z) = \sum_{n=1}^{\infty} n P_{1n} z^{n-1} \quad (3.8)$$

$$I_2(z) = \sum_{n=1}^{\infty} P_{2n} z^n, \quad I'_2(z) = \sum_{n=1}^{\infty} n P_{2n} z^{n-1} \quad (3.9)$$

Such that $I_0(1) + I_1(1) + I_2(1) = 1$.

On simplifying equation (3.6) by using equation (3.5) we get,

$$I_0(z) = \frac{1}{m\beta} [I_2(z)\{\lambda + \psi - \lambda z\} - (\lambda + \psi)zP_{21} + m\beta P_{00} + m\beta zP_{01}] \quad (3.10)$$

Taking $\lim_{z \rightarrow 1} I_0(z)$, we get

$$I_0(1) = \frac{1}{m\beta} [\psi I_2(1) - (\lambda + \psi)P_{21} + m\beta P_{00} + m\beta P_{01}] \quad (3.11)$$

Where $I_0(1) = P(WV) =$ Probability of customers serving in state of working vacation by server.

On simplifying equation (3.2) by using equation (3.1), we get,

$$\begin{aligned} \frac{1}{m\beta} [\lambda z + \eta z + \theta p z + \beta m z - \lambda z^2 - \eta] [I_2(z)(\lambda + \psi - \lambda z) - (\lambda + \psi)zP_{21} + \\ m\beta P_{00} + m\beta zP_{01}] = (\lambda z + \eta z + \theta p z + m\beta z - \eta)P_{00} - \eta z P_{01} - \alpha z P_{01} + \alpha z(1 - z)I'_0(z) \end{aligned} \quad (3.12)$$

Taking limit z tends to 1 in (3.12), we get

$$I_2(1) = \frac{1}{\psi} \left[(\lambda + \psi)P_{21} + \frac{m\beta}{(m\beta + \theta p)} \{ \lambda P_{00} - (m\beta + \theta p + \eta + \alpha)P_{01} \} \right] \quad (3.13)$$

where $I_2(1) = P(C) =$ Probability that server is in complete vacation.

On simplifying equation (3.4) by using equation (3.3), we get

$$\begin{aligned} I_1(z) \{ d\lambda z + s\mu z - d\lambda z^2 - \mu s \} = \theta p z I_0(z) + \psi z I_2(z) - \theta p z P_{00} - \theta p z^2 P_{01} - \psi z^2 P_{21} + \\ P_{11} \{ (d\lambda + s\mu)z^2 - s\mu z \} - s\mu z^2 P_{12} \end{aligned} \quad (3.14)$$

Taking limit z tends to 1 in (3.14), we get

$$I_1(1) = \frac{1}{(d\lambda - s\mu)} [\psi I_2(1) + \theta p I_0(1) - \theta p P_{00} - \mu(1 + s)P_{11}] \quad (3.15)$$

Where, $I_1(1) = P(B) =$ Probability of serving customer by server in busy state.

Now on differentiating equation (3.10) with respect to z and substituting limit z tends to 1 in it, we get

$$I'_0(1) = \frac{1}{(\gamma p + m\phi)} [I_1(1)(s\mu - d\lambda) - \theta p I_0(1) - (\lambda + \psi)I_2(1) - (\lambda + \psi)P_{21} + (m\beta + 2\theta p)P_{01} + \theta p P_{00} + 2\psi P_{21} - (2d\lambda + s\mu)P_{11} + 2s\mu P_{12}] \quad (3.16)$$

Where $I'_0(1) = E[L_0]$ = Expected queue length during working vacation,

Now differentiate equation (3.14) w.r.t. z and substituting limit z tend to 1 in it, we get

$$I'_1(1) = \frac{1}{(r\lambda - s\mu)} [\psi I'_2(1) + \theta p I'_0(1) - \mu P_{11} - s\mu I_1(1)] \quad (3.17)$$

Where $I'_1(1) = E[L_1]$ = Expected queue length during busy state,

Now differentiate equation (3.10) w.r.t. z and substituting limit z tends to 1 in it, we get

$$I'_2(1) = \frac{1}{\psi} [\lambda I_2(1) + m\beta I'_0(1) + (\lambda + \psi)P_{21} - m\beta P_{01}] \quad (3.18)$$

Where $I'_2(1) = E[L_2]$ = Expected queue length when server is in complete vacation.

Now, Expected queue length of system is $E[L] = E[L_0] + E[L_1] + E[L_2]$ and Abandonment rate of a customer in WV state is $R = \xi E[L_0]$.

Rearrange terms of recurrence relation (3.1), (3.2), (3.3), (3.4), (3.5), (3.6) we get,

$$\begin{aligned} P_{21} &= \left[\frac{m\beta}{(\lambda + \psi)} \right] P_{01} = K_1 P_{01}, \text{ where } K_1 = \left[\frac{m\beta}{(\lambda + \psi)} \right] \\ P_{00} &= \left(\frac{\eta + \xi}{\lambda} \right) P_{01} = K_2 P_{01}, K_2 = \left(\frac{\eta + \xi}{\lambda} \right) \\ P_{11} &= \frac{1}{\mu} \left(\frac{m\beta\psi}{(\lambda + \psi)} + \theta p \right) P_{01} = K_3 P_{01}, K_3 = \frac{1}{\mu} \left(\frac{m\beta\psi}{(\lambda + \psi)} + \theta p \right) \\ P_{12} &= \left(\frac{d\lambda}{s\mu} \right) P_{11} = \left(\frac{d\lambda}{s\mu} \right) K_3 P_{01} = K_4 P_{01}, K_4 = \left(\frac{d\lambda}{s\mu} \right) K_3 \end{aligned}$$

By using above P_{nj} 's, we can rewrite $I_0(1), I_1(1), I_2(1)$ in terms of P_{01} as follows,

$$\begin{aligned} I_2(1) &= T_2 P_{01}, \text{ where } T_2 = \frac{1}{\psi} \left[(\lambda + \psi)K_1 + \frac{m\beta}{(m\beta + \theta p)} \{ \lambda K_2 - (m\beta + \theta p + \eta + \alpha) \} \right] \\ I_0(1) &= T_0 P_{01}, \text{ where } T_0 = \frac{1}{m\beta} [\psi T_2 - (\lambda + \psi)K_1 + m\beta (K_2 + 1)] \\ I_1(1) &= T_1 P_{01}, \text{ where } T_1 = \frac{1}{(d\lambda - s\mu)} [\psi T_2 + \theta p T_0 - \theta p K_2 - \mu(1 + s)K_3] \end{aligned}$$

Since, $I_0(1), I_1(1), I_2(1), I'_0(1), I'_1(1), I'_2(1), P_{21}, P_{00}, P_{11}, P_{12}$ all are expressed in terms of P_{01} . Therefore, By normalizing condition

$I_0(1) + I_1(1) + I_2(1) = 1$, we can easily obtain the value of P_{01} as
 $\therefore P_{01} = \frac{1}{[T_0 + T_1 + T_2]} = [T_0 + T_1 + T_2]^{-1}$

4 Graphical Illustration

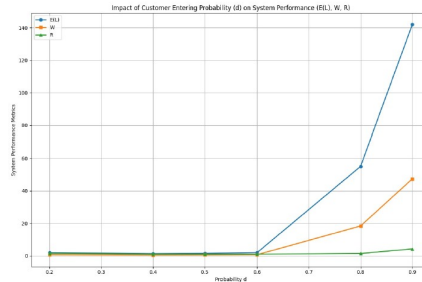


Figure 2: Impact of d on $E(L)$, W , R .

Figure 2 shows effect of customers entering probability d in busy state on various system performances $E(L), W, R$. It is clear from Figure 2, if we take the parameters $\lambda = 3, \mu = 5, \eta = 4, \alpha = 1, \psi = 1, \beta = 0.5, \theta = 0.5, p = 0.5, m = 0.5, s = 0.5$ and increase the value of entering probability d i.e. as d increases R remains almost constant for starting values of d and then gradual increment has been noticed while for $E(L)$ and W , as d increases, starting with a slow rise abruptly goes high.

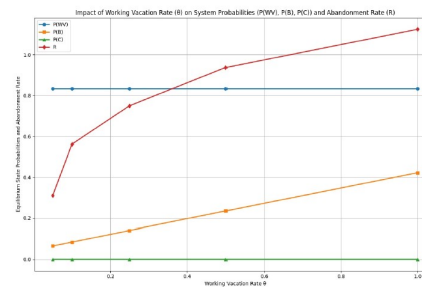


Figure 3: Impact of θ on $P(B), P(WV), P(C)$ and R .

Figure 3 shows effect of working vacation rate θ on various equilibrium or system probabilities and abandonment rate R of customers in system. It is clear from Figure 3, if we take the parameters $\lambda = 3, \mu = 5, \eta = 4, \alpha = 1, \psi = 1, \beta = 0.5, d = 0.5, p = 0.5, m = 0.5, s = 0.5$ and as the value of working vacation rate θ increases, $P(B)$ and R increases while $P(WV)$ and $P(C)$ remains constant.

5 Application of the Model

This model captures the flexibility required in real-world systems (Healthcare system, banking services, IT service system, customer service call center, etc.) that must dynamically adjust to client needs, priority variances, and resource availability. By implementing a structured system like this, service providers can improve efficiency, balance demand with operational capacity, and reduce client drop-off rates.

6 Conclusion

This study focuses on customers shifting priorities under a Bernoulli scheduling policy during busy state, customer impatience in working vacations, complete vacation in a M/M/1 system. Key performance metrics, including expected queue lengths in different states (busy, working vacation, and complete vacation), average sojourn time, and abandonment rates during the WV phase, were derived using the PGF method. The impact of specific parameters on system metrics demonstrated through graphical analysis. The findings suggest that, in many service-based industries, customers frequently adjust their priorities according to personal needs or situational changes. Service providers must actively address these shifting customer needs to prevent client loss. Monitoring customer behavior and adapting service delivery based on urgency and other factors can significantly improve the customer experience. Additionally, in queuing models that incorporate working vacations, a reduced service rate may increase customer impatience, especially if wait times are prolonged. Effective management strategies are crucial for minimizing the adverse effects on customers patience during these slower service periods.

Future Scope: This model could be further expanded for non-Markovian system by incorporating multi servers, offering additional flexibility in managing customer needs.

References

- [1] Ahmed, F.I., Alemu, S.D., & Tilahu, G.T. (2021). A single server markovian queue with single working vacation, state dependent reneging and retention of reneged customers. *Ethiopian Journal of Education and Sciences*, 17(1), 51–69.
- [2] Altman, E., & Yechiali, U. (2006). Analysis of customers' impatience in queues with server vacations. *Queueing Systems: Theory and Applications*, 52(4), 261–279.
- [3] Ancker, J. C.J., & Gafarian, A.V. (1963). Some queuing problems with balking and reneging. I. *Operations Research*, 11(1), 88–100.
- [4] Ancker, J. C.J., & Gafarian, A.V. (1963). Some queuing problems with balking and reneging-II. *Operations Research*, 11(6), 928–937.
- [5] Choudhury, G., & Paul, M. (2006). A two-phase queueing system with bernoulli vacation schedule under multiple vacation policy. *Statistical Methodology*, 3(2), 174–185.
- [6] Devi, V. N.R., Rao, A.A., & Chandan, K. (2019). M/M/1 queue with working vacation, server failure and customer's impatience. *International Journal of Scientific & Technology Research*, 8(09), 2262–2268.
- [7] Goswami, V. (2014). Analysis of impatient customers in queues with bernoulli schedule working vacations and vacation interruption. *Journal of Stochastics*, 1–10.
- [8] Gupta, P., & Kumar, N. (2021). Cost optimization of single server retrial queueing model with bernoulli schedule working vacation, vacation interruption, and balking. *Journal of Mathematical and Computational Science*, 11(3), 2508–2523.
- [9] Haight, F.A. (1957). Queueing with balking. *Biometrika*, 44(3-4), 360–369.
- [10] Haight, F.A. (1959). Queueing with reneging. *Metrika*, 2, 186–197.
- [11] Keilson, J., & Servi, L.D. (1986). Oscillatory random walk models for GI/G/1 vacation systems with bernoulli schedules. *Journal of Applied Probability*, 23(3), 790–802.

- [12] Kella, O. (1990). Optimal control of the vacation scheme in an M/G/1 queue. *Operations Research*, 38(4), 724–728.
- [13] Laxmi, P.V., & Jyothsna, K. (2014). Performance analysis of variant working vacation queue with balking and reneging. *International Journal of Mathematics in Operational Research*, 6(4), 505.
- [14] Li, W., Shi, D., & Chao, X. (1997). Reliability analysis of M/G/1 queueing systems with server breakdowns and vacations. *Journal of Applied Probability*, 34(2), 546–555.
- [15] Lin, C.-H., & Ke, J.C. (2009). Multi-server system with single working vacation. *Applied Mathematical Modelling*, 33(7), 2967–2977.
- [16] Liu, W.Y., Xu, X.L., & Tian, N.-S. (2007). Stochastic decompositions in the M/M/1 queue with working vacations. *Operations Research Letters*, 35(5), 595–600.
- [17] Manoharan, P., & Jeeva, T. (2020). Impatient customers in a markovian queue with bernoulli schedule working vacation interruption and setup time. *Applications and Applied Mathematics: An International Journal (AAM)*, 15(2), 725,739.
- [18] Perel, N., & Yechiali, U. (2010). Queues with slow servers and impatient customers. *European Journal of Operational Research*, 201(1), 247–258.
- [19] Rathore, R. (2022). A review on study of application of queueing models in hospital sector. *International Journal for Global Academic & Scientific Research*, 1(2), 1–6.
- [20] Rathore, R. (2022). A study on application of stochastic queueing models for control of congestion and crowding. *International Journal for Global Academic & Scientific Research*, 1(1).
- [21] Rathore, R., & Shrivastava, R.K. (2024). Analysis of an M/M/1/o model with bernoulli schedule during busy period, additional task, customers fluctuating priorities behavior during working vacation. *International Journal of Statistics and Applied Mathematics*, 9(5), 129–136.
- [22] Selvaraju, N., & Goswami, C. (2013). Impatient customers in an M/M/1 queue with single and multiple working vacations. *Computers & Industrial Engineering*, 65(2), 207–215.
- [23] Servi, L.D., & Finn, S.G. (2002). M/M/1 queues with working vacations (M/M/1/WV). *Performance Evaluation*, 50(1), 41–52.
- [24] Shrivastava, R.K., & Rathore, R. (2024). Analysis of single server markovian queueing model with differentiated working vacation, vacation interruption, soft failure, reneging of customers. *International Journal for Global Academic & Scientific Research*, 3(3), 1–13.
- [25] Swathi, C., & Vasanta Kumar, V. (2018). Analysis of M/M/1 queueing system with customer reneging during server vacations subject to server breakdown and delayed repair. *International Journal of Engineering & Technology*, 7(4.10), 552–557.
- [26] Tian, N., Zhao, X., & Wang, K. (2008). The M/M/1 queue with a single working vacation. *International Journal of Information and Management Sciences*, 19(4), 621–634.
- [27] Yang, D. Y., & Wu, Y. Y. (2017). Analysis of a finite-capacity system with working breakdowns and retention of impatient customers. *Journal of Manufacturing Systems*, 4, 207–216.
- [28] Yechiali, U. (2007). Queues with system disasters and impatient customers when the system is down. *Queueing Systems: Theory and Applications*, 56(3-4), 195–202.

- [29] Yue, D., Yue, W., & Xu, G. (2012). Analysis of customers' impatience in an M/M/1 queue with working vacations. *Journal of Industrial and Management Optimization*, 8(4), 895–908.
- [30] Zhang, H., & Shi, D. (2009). The M/M/1 queue with bernoulli schedulecontrolled vacation and vacation interruption. *International Journal of Information and Management Sciences*, 20(4), 579–587.