Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 11, Number 4 (2015), pp. 2561-2579

http://www.ripublication.com

Optimal Control of an N-policy two-phase M/M/1Gated Queueing System with Server startup, Breakdowns, Delayed Repair and balking

V. Vasanta Kumar[1], T.srinivasa Rao[2], B.Srinivasa Kumar[3]

Email id-vvkumar@kluniversity.in[1],tsr_2505@kluniversity.in[2], sk_bhavirisetty@kluniversity.in[3]
K. L. University, Vaddeswaram, Guntur(Dist.), Andhra Pradesh, INDIA.

Abstract

This paper investigates the economic behavior of the M/M/1 gated queue with server startup, customer balking, two-phases of compulsory service and breakdown period which consists of a delay period and repair period during second phase of service under N-policy. The first phase of service is a batch service to all waiting customers in the queue and the second phase of service is to each customer in the batch in FIFO order. The server is turned off each time the system empties. As and when the queue length reaches or exceeds N (threshold) batch service starts. Before the batch service, the system requires random startup time for pre-service. As soon as the startup period is over the server starts the batch service followed by individual service to all waiting customers in the batch queue. Arrivals during the batch service are not allowed to join the same batch which is in the first phase of service. The service station may breakdown during second phase of service and there may be delay in repair due to non-availability of the repairing facility.

Keywords: two-phase, vacation, breakdowns, N-policy, gating, repair time, delay time, cost function, balking.

1. Introduction

The service station in many queueing systems (eg: in communication systems, computer networks, flexible manufacturing systems etc) is subject to unpredictable breakdowns and can be repaired. Therefore, it is necessary to see how the breakdowns affect the level of performance of the system.

One may refer B. Srinivasa Kumar et al. [1] for the concise review of literature regarding the two-phase gated queueing systems with server start up, N-policy, breakdowns and delayed repair.

To our knowledge, to date, none of the studies has investigated the two-phase M/M/1 gated queue with N-policy, sever breakdowns and delayed repair with customer balking.

Thus the present study is aimed at the study of the economic behavior of an N-policy M/M/1 gated queue with service is in two phases and the server is typically subject to unpredictable breakdowns, customer balking and delay in repair . Arrival of customers is assumed to follow a Poisson process and waiting customers receive batch service all at a time in the first phase and proceed to the second phase to receive the individual service. The server is turned off each time the system empties.When the total number of customers in the system reaches the threshold N(N≥1), the server is immediately turned on but is temporarily unavailable to serve the waiting customers. The server requires time for the preparatory work (ie,. begin startup) before starting service. Once the startup is terminated the server immediately starts serving the waiting customers. By gating we mean that the customers who arrive during the batch service are not allowed to enter the batch which is already in service, but are served in the next visit of the server to the batch queue. The service station may breakdown when the server is in individual service and there may be delay in repair due to non-availability of the repairing facility. It is assumed that the batch service times, the repair times and the startup times obey exponential distribution.

2. Model description and assumptions

We consider the M/M/1 gated queueing system with server startup, two phases of service, unreliable server, customer balking and delay in repair, where the unreliable server operates under N-policy.

Assumptions of the model:

1. Customers are assumed to arrive according to Poisson process with mean arrival rate λ and join the batch queue. Arrived customers are admitted to service on a first come first served basis.
2. Upon arrival, the customers may decide not to join the queue at all due to a long queue. Let b_1 be the probability that the arriving customer join in the queue when the server is on vacation or in startup and b_2 be the probability that the arriving customer join the queue when the server is in batch service or in individual service or in break down or in repair states.
3. The service is in two phases. The first phase of service is batch service to all customers waiting in the queue. On completion of batch service, the server proceeds to the second phase to serve all customers in the batch individually. Batch service time is assumed to be exponentially distributed with mean $1/\beta$ and is independent of batch size. Individual service times are exponentially distributed with mean $1/\mu$. On completion of individual service, the server

returns to the batch queue to serve the customers who have arrived. If the customers are waiting, the server starts the batch service followed by individual service to each customer in the batch. If no customer is waiting the server takes a vacation.

4. Whenever the system becomes empty, the server is turned off. As soon as the total number of arrivals in the queue reaches or exceeds pre determined threshold N, the server is turned on and is temporarily unavailable for the waiting customers. The server needs a startup time which follows an exponential distribution with mean $1/\theta$. As soon as the server finishes startup, it starts serving the first phase of waiting customers.
5. The customers who arrive during the batch service are not allowed to join the batch which is in service and are served along with other arrivals during the next visit of the server to the batch queue.
6. The breakdowns are generated by an exogenous Poisson process with rate α. As soon as the breakdown of the server occurs, it is sent for repair during which the server stops providing service to the arriving customers and waits for repair to start, which we may refer to as waiting period of the server. The waiting time is defined as delay time and is assumed to be exponentially distributed with mean $1/\delta$ in any phase of service. Repair time is also exponentially distributed with mean $1/\gamma$.
7. Immediately after the server is repaired, it starts to serve and the service time is cumulative. A customer who arrives and finds the server busy or broken down will wait in the queue until the server is available. Customers continue to arrive during the delay and repair periods of the broken server.

The further discussions of the paper are organized as follows:
Section 2 deals with model descriptions and assumptions. Section 3 deals with development of the model. Expected number of customers at different states of the server is presented in Section 4. Characteristic futures of the system are presented in Section 5. Waiting time in the queue is presented in Section 6 through heuristic approach. Section 7 deals with Optimal control policy; sensitivity analysis is presented in Section 8. Numerical examples are provided in Section 9.

3. Development of the model.

In steady state the following notations are used.

$P_{0,0,0}$=The probability that there is no customer in the system and the server is on vacation.

$P_{0,i,0}$=The probability that there are i customers waiting in the batch queue and the server is on vacation, i=1, 2, 3,.., N-1.

$P_{1,i,0}$=The probability that there are i customers in the batch queue and the server is doing pre-service (startup work), where i=N, N+1, N+2,…

$P_{2,i,0}$ =The probability that there are i customers in the batch which is in batch service, i=1, 2, 3,…

$p_{3,i\,j}$ =The probability that there are i customers in the batch queue and j customers in the individual queue and the server is in individual service, i=0,1,2, 3,… and j=1,2,3,…

$P_{4,i,j}$ =The probability that there are i customers in the batch queue and j customers in the individual queue while the server is in individual service, but found to be broken down and waiting for repair, i=0, 1, 2, 3,… and j=1,2,3,…

$P_{5,i,j}$ =The probability that there are i customers in the batch queue and j customers in the individual queue while the server is in individual service but is under repair, i=0, 1, 2,3… and j=1, 2, 3,…

The steady state equations satisfied by the system size probabilities are as follows:

$$\lambda b_1 P_{0,0,0} = \mu P_{3,0,1}. \tag{1}$$

$$\lambda b_1 P_{0,i,0} = \lambda b_1 P_{0,i-1,0},\ 1 \le i \le N\text{-}1. \tag{2}$$

$$(\lambda b_1 + \theta) P_{1,N,0} = \lambda b_1 P_{0,N-1,0}. \tag{3}$$

$$(\lambda b_1 + \theta) P_{1,i,0} = \lambda b_1 P_{1,i-1,0},\ i > N. \tag{4}$$

$$\beta P_{2,i,0} = \mu P_{3,i,1},\ 1 \le i \le N\text{-}1. \tag{5}$$

$$\beta P_{2,i,0} = \mu P_{3,i,1} + \theta P_{1,i,0},\ i \ge N. \tag{6}$$

$$(\lambda b_2 + \alpha + \mu) P_{3,0,j} = \mu P_{3,0,j+1} + \pi_0 \beta P_{2,j,0} + \gamma P_{5,0,j},\ j \ge 1. \tag{7}$$

$$(\lambda b_2 + \alpha + \mu) P_{3,i,j} = \mu P_{3,i,j+1} + \pi_i \beta P_{2,j,0} + \lambda b_2 P_{3,i-1,j} + \gamma P_{5,i,j},\ i \ge 1, j \ge 1. \tag{8}$$

$$(\lambda b_2 + \delta) P_{4,0,j} = \alpha P_{3,0,j},\ j \ge 1. \tag{9}$$

$$(\lambda b_2 + \delta) P_{4,i,j} = \alpha P_{3,i,j} + \lambda b_2 P_{4,i-1,j},\ i \ge 1, j \ge 1. \tag{10}$$

$$(\lambda b_2 + \gamma) P_{5,0,j} = \delta P_{4,0,j},\ j \ge 1. \tag{11}$$

$$(\lambda b_2 + \gamma) P_{5,i,j} = \delta P_{4,i,j} + \lambda b_2 P_{5,i-1,j},\ i \ge 1, j \ge 1. \tag{12}$$

where $\pi_i = \dfrac{(\lambda b_2)^i \beta}{(\lambda b_2 + \beta)^{i+1}}$.

To determine the probability distribution of the number of customers in the system and hence the expected number of customers in the system, the following probability generating functions are defined.

$$G_0(z) = \sum_{i=0}^{N-1} P_{0,i,0} z^i , \qquad G_1(z) = \sum_{i=N}^{\infty} P_{1,i,0}\, z^i , \qquad G_2(z) = \sum_{i=1}^{\infty} P_{2,i,0}\, z^i ,$$

$$G_3(z,y) = \sum_{i=0}^{\infty}\sum_{j=1}^{\infty} P_{3,i,j}\, z^i y^j , \; G_4(z,y) = \sum_{i=0}^{\infty}\sum_{j=1}^{\infty} P_{4,i,j}\, z^i y^j , \; G_5(z,y) = \sum_{i=0}^{\infty}\sum_{j=1}^{\infty} P_{5,i,j}\, z^i y^j$$

$$R_j(z) = \sum_{i=0}^{\infty} P_{3,i,j}\, z^i , S_j(z) = \sum_{i=0}^{\infty} P_{4,i,j} z^i , T_j(z) = \sum_{i=0}^{\infty} P_{5,i,j} z^i \text{ and } \Pi(z) = \sum_{i=0}^{\infty} \pi_i\, z^i = \frac{\beta}{(\lambda b_2 (1-z) + \beta)},$$

where $|z| \le 1$ and $|y| \le 1$.

$$(1-z)\, G_0(z) = (1-z^N)\, P_{0,0,0}. \tag{13}$$

$$[\, \lambda b_1 (1-z) + \theta \,] G_1(z) = \lambda\ b_1\, z^N\, P_{0,0,0}. \tag{14}$$

$$\beta\, G_2(z) = \mu\, R_1(z) + \theta\, G_1(z) - \lambda b_1 P_{0,0,0}. \tag{15}$$

$$[\, \lambda\, b_2 (1-z) + \alpha + \mu \,] R_j(z) = \mu\, R_{j+1}(z) + \gamma\, T_j(z) + \beta\, P_{2,j,0}\, \pi(z) . \tag{16}$$

$$\left[\lambda\, y\, b_2 (1-z) + \alpha\, y + \mu (y-1)\right] G_3(z,y) =$$

$$\beta\, y\ \pi(z)\ G_2(y) + \gamma\, y\, G_5(z,y) - \mu\, y\, R_1(z). \tag{17}$$

$$\left[\lambda b_2 (1-z) + \delta\right] S_j(z) = \alpha\, R_j(z). \tag{18}$$

$$\left[\lambda\, b_2 (1-z) + \delta\right] G_4(z,y) = \alpha G_3(z,y). \tag{19}$$

$$\left[\lambda\, b_2 (1-z) + \delta\right] S_j(z) = \alpha\, R_j(z). \tag{20}$$

$$\left[\lambda\, b_2 (1-z) + \gamma\right] G_5(z,y) = \delta\, G_4(z,y). \tag{21}$$

The total probability generating function $G(z,y)$ is given by

$$G(z,y) = G_0(z) + G_1(z) + G_2(z) + G_3(z,y) + G_4(z,y) + G_5(z,y). \tag{22}$$

The normalizing condition is

$$G(1,1) = G_0(1) + G_1(1) + G_2(1) + G_3(1,1) + G_4(1,1) + G_5(1,1) = 1. \tag{23}$$

From equations (13) to (21),we get

$$G_0(1) = NP_{0,0,0}, \tag{24}$$

$$G_1(1) = (\lambda b_1/\theta)P_{0,0,0}, \tag{25}$$

$$G_2(1) = \mu R_1(1)/\beta, \tag{26}$$

$$G_3(1,1) = \frac{\left[\frac{\lambda b_2 \mu R_1(1)}{\beta} + \theta G_1{}^1(1)\right]}{\mu\left(1 - \frac{\lambda b_2}{\mu}(1 + \frac{\alpha}{\gamma} + \frac{\alpha}{\delta})\right)}, \tag{27}$$

$$G_4(1,1) = \frac{\alpha}{\delta} G_3(1,1) \text{ and} \tag{28}$$

$$G_5(1,1) = \frac{\alpha}{\gamma} G_3(1,1). \tag{29}$$

The normalizing condition (6.23) yields,

$$R_1(1) = \frac{\lambda b_2}{\mu} + \frac{(b_2 - b_1)}{b_2}\frac{\beta \rho_1 (1-\rho_1)}{\mu} - \frac{\lambda \rho_1}{\mu}(b_2 - b_1),$$

where $\rho_1 = \frac{\lambda b_2}{\mu}(1 + \frac{\alpha}{\gamma} + \frac{\alpha}{\delta})$.

Probability that the server is neither in batch service nor in individual service is given by

$G_0(1) + G_1(1) = 1 - \frac{\lambda b_2}{\beta} - \frac{\lambda b_2}{\mu}(1 + \frac{\alpha}{\gamma} + \frac{\alpha}{\delta})$.This gives

$$P_{0,0,0} = (1-\rho)\frac{\theta}{(\lambda b_1 + N\theta)}, \tag{30}$$

where $\rho = \frac{\lambda b_2}{\beta} + \frac{\lambda b_2}{\mu}(1 + \frac{\alpha}{\gamma} + \frac{\alpha}{\delta})$ is the utilizing factor of the system.

From Equation (30) we have $\rho < 1$, which is the necessary and sufficient condition under which steady state solution exits.
Using the condition $\underset{y\to 1}{Lt}\, G_3(1,y) = \underset{z\to 1}{Lt}\, G_3(z,1)$, we obtain

$$R_1^1(1) = \frac{\lambda b_2}{\beta(1-\rho_1)} R_1(1) + \frac{\lambda b_1 \rho_1}{\mu} - \frac{\lambda^2 b_1 b_2 \rho_1}{\mu\beta(1-\rho_1)}$$

Under steady state, let $P_v, P_s, P_b, P_i, P_{bi}$ and P_{di} be the probabilities that the server is in vacation, in startup, in batch service, in individual service, waiting for repair during individual service and under repair during individual service states, respectively. Then

$$P_v = G_0(1) = N\, P_{0,0,0}, \tag{31}$$

$$P_s = G_1(1) = (\lambda b_1 / \theta).P_{0,0,0}, \tag{32}$$

$$P_b = G_2(1) = (\mu R_1(1) / \beta), \tag{33}$$

$$P_i = G_3(1,1) = \frac{\left[\frac{\lambda b_2\ \mu R_1(1)}{\beta} + \theta G_1^1(1)\right]}{\mu\left(1 - \frac{\lambda b_2}{\mu}(1 + \frac{\alpha}{\gamma} + \frac{\alpha}{\delta})\right)}, \tag{34}$$

$$P_{bi} = G_4(1,1) = \frac{\alpha}{\delta} G_3(1,1) \quad \text{and} \tag{35}$$

$$P_{di} = G_5(1,1) = \frac{\alpha}{\gamma} G_3(1,1), \tag{36}$$

where $G_1^1(1) = \lambda b_1 \left(\frac{\lambda b_1 + N\theta}{\theta^2}\right) P_{0,0,0}$.

4. Expected number of customers at different states of the server

Using the probability generating functions expected number of customers in the system at different states are presented in this section
Let $L_v, L_s, L_b, L_i, L_{bi}$ and L_{di} be the expected number of customers in the system when the server is in vacation, in startup, in batch service, waiting for repair during

batch service, in individual service, waiting for repair during individual service and under repair during individual service states, respectively.
Then,

$$L_v = \sum_{i=1}^{N-1} i\, P_{0,i,0} = G_0^{1}(1) = \frac{N(N-1)}{2} P_{0,0,0}. \tag{37}$$

$$L_s = \sum_{i=1}^{\infty} i\, P_{1,i,0} = G_1^{1}(1) = \frac{\lambda b_1 (\lambda b_1 + N\theta)}{\theta^2} P_{0,0,0}. \tag{38}$$

$$L_b = \sum_{i=1}^{\infty} i\, P_{2,i,0} = G_2^{1}(1) = (\mu / \beta)\, R_1^{1}(1) + (\theta / \beta)\, G_1^{1}(1). \tag{39}$$

$$L_i = \sum_{i=0}^{\infty}\sum_{j=1}^{\infty} (i+j)\, P_{3,i,j} = G_3^{1}(1,1).$$

$$= \left[\frac{\rho_1}{(1-\rho_1)} + \frac{\lambda b_2}{2(1-\rho_1)}\left(1-\frac{\lambda b_2}{\mu}\right)\left(\frac{1}{\gamma}+\frac{1}{\delta}\right) - \frac{\lambda^2 b_2^{2} \alpha}{\gamma \mu \delta (1-\rho_1)} \right] G_3(1,1) +$$

$$\frac{\lambda b_1}{\mu (1-\rho_1)} \left[\frac{N(N-1)}{2} P_{0,0,0} + \left(\frac{\lambda b_1}{\theta} + \frac{\lambda b_2}{\beta} + 1 \right) \left(\frac{\lambda b_1 + N\theta}{\theta} \right) P_{0,0,0} + \frac{b_2 \mu}{b_1 \beta} \left(R_1^{1}(1) + R_1(1) \right) \frac{\lambda b_2^{2}}{b_1 \beta^2} \mu \; R_1(1) \right]$$

$$- \frac{\lambda b_2}{\mu(1-\rho_1)} \left(\frac{1}{\gamma} + \frac{1}{\delta} \right) \left[\lambda b_1 \left(\frac{\lambda b_1 + N\theta}{\theta} \right) P_{0,0,0} + \frac{\lambda b_2 \mu R_1(1)}{\beta} \right] \tag{40}$$

$$L_{bi} = \sum_{i=0}^{\infty}\sum_{j=1}^{\infty} (i+j) P_{4,i,j} = G_4^{1}(1,1) = \frac{\lambda \alpha b_2}{\delta^2} G_3(1,1) + \frac{\alpha}{\delta} G_3^{1}(1,1) \; . \tag{41}$$

$$L_{di} = \sum_{i=0}^{\infty}\sum_{j=1}^{\infty} (i+j) P_{5,i,j} = G_5^{1}(1,1) = \frac{\lambda \alpha b_2}{\gamma} \left(\frac{1}{\gamma} + \frac{1}{\delta}\right) G_3(1,1) + \frac{\delta}{\gamma} G_3^{1}(1,1) \; . \tag{42}$$

Let $G_n^{1}(.)$ denotes the first order derivative of $G_n(.)$.

The expected number of customers in the system is given by
$L(N) = L_v + L_s + L_b + L_i + L_{bi} + L_{di}$

$$= \frac{N(N-1)}{2} P_{0,0,0} + \frac{\lambda b_1 (\lambda b_1 + N\theta)}{\theta^2} P_{0,0,0} + \frac{\mu R_1^{1}(1) + \theta G_1^{1}(1)}{\beta}$$

$$+\left(1+\frac{\alpha}{\delta}+\frac{\alpha}{\gamma}\right) G_3^{1}(1,1) + \left(\frac{\lambda \alpha b_2}{\delta^2} + \frac{\lambda \alpha b_2}{\delta \gamma} + \frac{\lambda \alpha b_2}{\gamma^2}\right) G_3(1,1). \qquad (43)$$

Special case
When $b_1=b_2=1$(i.e all the arriving customers join the queue), equation(43) becomes

$$L(N) = \frac{\rho_1}{(1-\rho_1)} + \frac{N(N-1)}{2(1-\rho_1)} P_{0,0,0} + \frac{\lambda(\lambda + N\theta)}{\theta^2 (1-\rho_1)} P_{0,0,0} + \frac{\lambda \rho_1^{2}}{(1-\rho_1)}\left(\frac{1}{\delta}+\frac{1}{\gamma}\right) + \frac{\lambda}{\beta(1-\rho_1)}$$

$$+\frac{\lambda^2 \alpha}{\mu \gamma \delta}\left(1+\frac{\gamma}{\delta}+\frac{\delta}{\gamma}\right) + \frac{\lambda^2 \rho_1}{\beta^2 (1-\rho_1)} - \frac{\lambda^2 \rho_1}{\mu(1-\rho_1)}\left(\frac{1}{\delta}+\frac{1}{\gamma}\right) - \frac{\lambda^2 \alpha \rho_1}{\mu \gamma \delta (1-\rho_1)} \qquad (44)$$

where $\rho_1 = \frac{\lambda}{\mu}(1+\frac{\alpha}{\gamma}+\frac{\alpha}{\delta})$, and $P_{0,0,0} = \left(1-\frac{\lambda}{\beta}+\frac{\lambda}{\mu}\left(1+\frac{\alpha}{\gamma}+\frac{\alpha}{\delta}\right)\right)\frac{\theta}{(\lambda + N\theta)}$.

This is the expected system length for the two-phase M/M/1 queueing system with server startup, N-policy, breakdowns and delayed repair.

5. Characteristic features of the system

In this section, we obtain the expected system length when the server is in different states. Let $E_{v,}\ E_s, E_b,\ E_i, E_{bi}$ and E_{di} denote the expected length of vacation period, startup period, batch service period, individual service period, delay period during individual service and repair period during individual service respectively. Then the expected length of a busy cycle is given by

$$E_c = E_v + E_s + E_b + E_i + E_{bi} + E_{di}. \qquad (45)$$

The long run fractions of time the server is in different states are respectively,

$$E_v / E_c = P_v = N P_{0,0,0}. \qquad (46)$$

$$E_s / E_c = P_s = (\lambda\, b_1\, P_{0,0,0}) / \theta. \tag{47}$$

$$E_b / E_c = P_b = \mu\, R_1(1) / \beta. \tag{48}$$

$$E_i / E_c = P_i = \frac{\lambda b_2\ \mu R_1(1) + \theta\, \beta\, G_1^{\,1}(1)}{\beta\, \mu \left(1 - \frac{\lambda\, b_2}{\mu} (1 + \frac{\alpha}{\gamma} + \frac{\alpha}{\delta}) \right)} \tag{49}$$

$$E_{bi} / E_c = P_{bi} = \frac{\alpha}{\delta} G_3(1,1) \tag{50}$$

$$E_{di} / E_c = P_{di} = \frac{\alpha}{\gamma} G_3(1,1) \tag{51}$$

Expected length of vacation period is given by

$$E_v = N / \lambda\, b_1 . \tag{52}$$

Substituting this in equation (46)

$$E_c = 1 / (\lambda\, b_1\, P_{0,0,0}) . \tag{53}$$

6. Heuristic Approach to waiting time in the queue.

An arbitrary customer waits different time amounts according to the state of his arriving epoch. First, we divide the regeneration cycle into six parts of the idle period, the startup period, the first phase batch service, the second phase individual service, the waiting time for repair due to breakdown in second phase and the repair time with respective probabilities

$$N\, P_{0,0,0},\ \frac{\lambda\ b_1\ P_{0,0,0}}{\theta}, \frac{\mu\, R_1(1)}{\beta},\ G_3(1,1),\quad \frac{\alpha}{\delta} G_3(1,1) \quad \text{and}\ \frac{\alpha}{\gamma} G_3(1,1).$$

That is the system states that the arriving customer sees the queue and determines his waiting time. The test customer has to wait during the individual service times for those already waiting (except the ongoing individual service) in the system. In addition to it,

(i) If the server is in idle state, the customer has to wait for the remaining idle period, startup period, the first phase batch service period.

(ii) If the server is in startup state, the customer has to wait for the remaining startup period, the first phase batch service period.

(iii) If the server is in the first phase service, the customer has to wait for the remaining time of the ongoing batch service and its individual service periods, plus the batch service period in which he gets service.

(iv) If the server is in the second phase service, the customer has to wait for the remaining time of the ongoing individual service plus the batch service.

(v) If the server is waiting for repair due to breakdown, the customer has to wait for the remaining breakdown period and the repair period plus the first phase batch service period.

(vi) If the server is in repair state due to breakdown, the customer has to wait for the remaining repair period and the first phase batch service period.

Thus

$$E(W_q) = L(N-1)\frac{1}{\mu} + \left(\frac{N-1}{2\lambda} + \frac{1}{\theta} + \frac{1}{\beta}\right) N P_{0,0,0} + \left(\frac{1}{\theta} + \frac{1}{\beta}\right)\frac{\lambda b_1 P_{0,0,0}}{\theta}$$

$$+\left(\frac{2}{\beta} + \frac{L(N)}{\mu}\right)\left(\frac{\mu R_1(1)}{\beta}\right) + \left(\frac{1}{\mu} + \frac{1}{\beta}\right) G_3(1,1) + \left(\frac{1}{\delta} + \frac{1}{\gamma} + \frac{1}{\beta}\right)\left(\frac{\alpha}{\delta}\right) G_3(1,1)$$

$$+\left(\frac{1}{\gamma} + \frac{1}{\beta}\right)\left(\frac{\alpha}{\gamma}\right) G_3(1,1) \qquad (54)$$

,

where $L(N) = L(N-1) + \rho$ and $L(N-1) = \lambda b_2 E(W_q)$. Substituting these values in the above expression

$$E(W_q) = \frac{1}{1-r}\left[\left(\frac{N-1}{2\lambda} + \frac{1}{\theta} + \frac{1}{\beta}\right) N P_{0,0,0} + \left(\frac{1}{\theta} + \frac{1}{\beta}\right)\frac{\lambda b_1}{\theta} P_{0,0,0} + \left(\frac{2\mu}{\beta} + \rho\right)\left(\frac{R_1(1)}{\beta}\right)\right.$$

$$\left. + \left(\frac{1}{\mu} + \frac{1}{\delta}(1+\frac{\alpha}{\delta})\right) + \frac{1}{\gamma}\left(\frac{\alpha}{\delta} + \frac{\alpha}{\gamma}\right) + \frac{1}{\beta}\left(\frac{\alpha}{\delta} + \frac{\alpha}{\gamma}\right)\right] G_3(1,1) \qquad (55)$$

where r= $\frac{\lambda b_2}{\mu} + \frac{\lambda b_2 R_1(1)}{\beta}$.

Reliability indices

Two reliability indices of the system viz. the system availability and failure frequency under steady state condition are given below. The system availability for a customer is the probability that the sever is working for customers, or in idle period or in startup period.

The steady state availability of the server will be given by

$G_0(1)+G_1(1)+G_2(1)+G_3(1,1)$

$$=\left(\frac{\lambda+N\,\theta}{\theta}\right)p_{0,0,0}+\frac{\mu\;R_1(1)}{\beta}+\frac{\left[\lambda\, b_2\;\mu\;R_1(1)\;+\;\theta\;\beta\;G_1^{1}(1)\right]}{\beta\;\mu\left(1-\dfrac{\lambda\, b_2}{\mu}\;(1+\dfrac{\alpha}{\gamma}+\dfrac{\alpha}{\delta})\right)} \quad (56)$$

The steady state failure frequency of the server is given by

$$F=\alpha\;G_3(1,1)=\frac{\alpha\left[\lambda\, b_2\;\mu\, R_1(1)+\theta\,\beta\, G_1^{\;1}(1)\right]}{\beta\,\mu\left(1-\dfrac{\lambda\, b_2}{\mu}(1+\dfrac{\alpha}{\gamma}+\dfrac{\alpha}{\delta})\right)} \quad (57)$$

7. Optimal control policy

In this section, we determine the optimal value of N that minimizes the long run average cost of two-phase M/M/1, N-policy queue with server break downs and delay in repair. To determine the optimal value of N we consider the following linear cost structure.

Let $T_A(N)$ be the average cost per unit of time. Then

$$T_A(N)=\;C_h\;\;L(N)\;\;+C_0\left(\frac{E_b}{E_c}+\frac{E_i}{E_c}\right)+C_m\left(\frac{E_s}{E_c}\right)+C_b\left(\frac{E_{bi}+E_{di}}{E_c}\right)+C_s\left(\frac{1}{E_c}\right)-C_r\left(\frac{E_v}{E_c}\right), \quad (58)$$

where

$C_h \equiv$ Holding cost per unit time for each customer present in the system,

$C_0 \equiv$ Cost per unit time for keeping the server on and in operation,

$C_m \equiv$ Start up cost per unit time,

$C_s \equiv$ Setup cost per cycle,

$C_b \equiv$ Break down cost per unit time for the unavailable server and

$C_r \equiv$ Reward per unit time as the server is doing secondary work in vacation.

From equations (49) to (52) it is observed that

$E_b\,/\,E_c\;,E_i\,/\,E_c\;,E_{bi}\,/\,E_c\;,\;E_{di}\,/\,E_c$ are not functions of the decision variable N.

Hence for determination of the optimal operating N-policy, minimizing $T_A(N)$ in equation (58) is equivalent to minimizing.

$$T_1(N)=\left[C_h\left[\frac{(b_1-b_2)\rho_1+b_2}{2\mu(1-\rho_1)}\right]\frac{N(N-1)}{2}+\frac{\lambda b_1}{\theta}C_m+\lambda b_1 C_s-N C_r\right].P_{0,0,0} \quad (59)$$

An approximate value of the optimal threshold N* can be found by solving the equation

$$\left.\frac{dT_1(N)}{dN}\right|_{N=N^*}=0$$

which is equivalent to

$$N^*=\frac{\lambda b_1}{\theta}+\sqrt{\frac{\lambda^2 b_1^2}{\theta^2}+\frac{\lambda b_1}{\theta}+\frac{\lambda b_1}{\theta}P(1-\lambda b_1 L)}, \quad (60)$$

where $P=\dfrac{C_r+C_m+\theta C_s}{KC_h}$ and $L=\dfrac{1}{\mu}\left(1+\dfrac{\alpha}{\gamma}+\dfrac{\alpha}{\delta}\right)$

It is noticed that $T_1^{11}(N^*)>0$ at this value of N* which suggests that $T_1(N)$ has only one local minimum and hence the optimal value N* is unique.
The exact value of the optimal threshold N* is the first N for which

$$T_1(N+1)\geq T_1(N). \quad (61)$$

8. **Sensitivity Analysis**

In this section, sensitivity analysis is performed on the optimum value N^* based on changes in the system parameters $\lambda, \mu, \beta, \theta, \alpha, \gamma$ and δ and the cost elements C_h, C_0, C_m, C_s, C_b and C_r. Differentiating N^* with respect to λ we obtain

$$\frac{\partial N^*}{\partial \lambda}=-\frac{b_1}{\theta}+\frac{b_1(2P+1)}{2\theta\sqrt{\frac{b_1\lambda}{\theta}(2P+1)}}, \text{ if } \frac{1-2PL\theta}{\theta^2}=0, \quad (62)$$

$$\frac{\partial N^*}{\partial \lambda}=-\frac{b_1}{\theta}+\frac{2(1-2PL\theta)\lambda b_1^2+b_1(2P+1)\theta}{2\theta\sqrt{(1-2PL\theta)b_1^2\lambda^2+(2P+1)\lambda b_1\theta}}, \text{ if } \frac{1-2kL\theta}{\theta^2}\neq 0. \quad (63)$$

where $P=\dfrac{C_r+\theta C_s+C_m}{C_h}$ and $L=\dfrac{1}{\mu}\left(1+\dfrac{\alpha}{\gamma}+\dfrac{\alpha}{\delta}\right)$.

Setting (62) and (63) to zero, and then solving for λ, we find

$$\lambda = \frac{(2\,b_1 P + 1)\,\theta}{4}, \text{ if } \frac{1 - 2\,P\,L\,\theta}{\theta^2} = 0 \qquad (64)$$

$$\lambda = \frac{(2\,P + 1)}{2\,b_1 \left(2PL - \sqrt{\dfrac{2\,P\,L}{\theta}} \right)}, \text{ if } \frac{1 - 2\,P\,L\,\theta}{\theta^2} \neq 0 \qquad (65)$$

Differentiating (62) and (63) with respect to N again and substituting the value of λ from (64) and (65) into the resulting differentiation respectively, we have

$$\frac{\partial^2 N^*}{\partial \lambda^2} = \frac{-2\,(2\,k + 1)^2}{\theta \left[\lambda\,(2\,k + 1)/\theta\right]^{3/2}} < 0, \text{ if } \frac{1 - 2\,k\,L\,\theta}{\theta^2} = 0, \qquad (66)$$

$$\frac{\partial^2 N^*}{\partial \lambda^2} = \frac{-2\left[2\,(1 - 2\,k\,\theta\,L) + (2\,k + 1)\,\theta\right]^2}{\theta\left[(1 - 2\,k\,\theta\,L)\,\lambda^2 + \lambda\,(2\,k + 1)\,\theta\right]^{3/2}} < 0, \text{ if } \frac{1 - 2\,k\,L\,\theta}{\theta^2} \neq 0. \qquad (67)$$

The above results show that the graph of N* is concave downward with respect to λ, which attains its maximum value under two parameter setting satisfying (64) and (65). Differentiating N* with respect to μ, we obtain

$$\frac{\partial N^*}{\partial \mu} = \frac{\lambda^2\, b_1^{\,2}\, P\,L}{\mu^2 \sqrt{\lambda^2 b_1^{\,2} + \lambda\, b_1\, \theta + 2\,\lambda\, b_1\, \theta\, P\,(1 - \rho_2)}} > 0. \qquad (68)$$

Thus, N^* is increasing in μ. Similarly differentiating N*with respect to α, γ and δ yields

$$\frac{\partial N^*}{\partial \alpha} = -\frac{P\, b_1^{\,2}\, \lambda^2}{\mu\, \gamma \sqrt{\lambda^2\, b_1^{\,2} + \lambda\, b_1\, \theta + 2\,\lambda\, b_1\, \theta\, P\,(1 - \rho_2)}} < 0, \qquad (69)$$

$$\frac{\partial N^*}{\partial \gamma} = \frac{P\,\lambda^2 b_1^{\,2}}{\mu\, \gamma^2 \sqrt{\lambda^2 b_1^{\,2} + \lambda\,\theta\, b_1 + 2\,\lambda\,\theta\, b_1\, P\,(1 - \rho_2)}} > 0, \qquad (70)$$

$$\frac{\partial N^*}{\partial \delta} = \frac{P\,\lambda^2\, b_1^{\,2}}{\mu\, \delta^2 \sqrt{\lambda^2\, b_1^{\,2} + \lambda\, b_1\, \theta + 2\,\lambda\,\theta\, b_1 P\,(1 - \rho_2)}} > 0, \qquad (71)$$

where $\rho_2 = \dfrac{\lambda b_1}{\mu}\left(1+\dfrac{\alpha}{\gamma}+\dfrac{\alpha}{\delta}\right)$

From (69), (70) and (71),we see that $N^*$is decreasing in α and is increasing in γ and δ.

9. Numerical examples

In this section, a numerical study was conducted to illustrate the model. For this, the cost parameters C_h=5, C_0=50, C_m=100, C_b=100, C_r=40, C_s=1000 and the following seven cases are considered.

Case1: We select μ=1.0,1.5, 2.0, 2.5, β=4, θ=3, α=0.05, γ =3,δ=3 and vary the values of λ.

Case2: We select λ=0.2, 0.4, 0.6, 0.8, β=4, θ=3, α=0.05, γ =3,δ=3 and vary the values of μ.

Numerical results of case1 and case2 are presented in Table1.

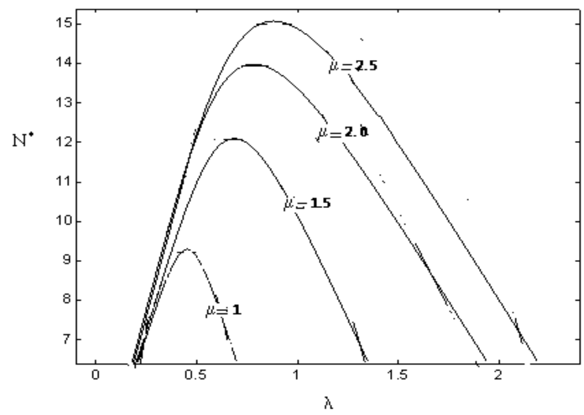

Fig.1. plots of (λ, N^*) with μ=1.0, 1.5, 2.0, 2.5, β=4, θ=3, α=0.05, γ=3, δ=3, C_h=5, C_0=50, C_m=100, C_b=100, C_r=40, C_s=1000

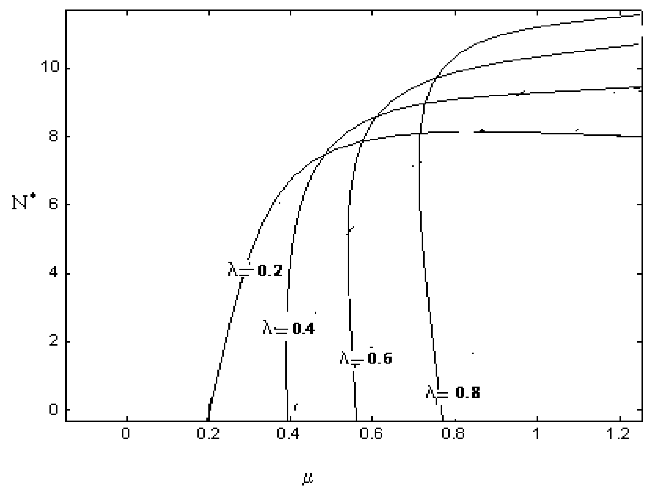

Fig.2. plots of (μ, N^*) with λ=0.2, 0.4, 0.6, 0.8, β=4, θ=3, α=0.05, γ =3, δ=3, C_h=5, C_0=50, C_m=100, C_b=100, C_r=40, C_s=1000

Table1. The optimal N^* and the minimum expected cost $T(N^*)$ with various (λ, μ)

β=4, θ=3, α=0.5, γ =3, δ=3, C_h=5, C_0=50, C_m=100, C_b=100, C_r=40, C_s=1000								
(λ,μ)	(0.2,1.5)	(0.4,1.5)	(0.6,1.5)	(0.8,1.5)	(0.3,0.5)	(0.3,1.0)	(0.3,1.5)	(0.3,2.0)
N^*:	8	10	10	9	4	8	9	9
$T(N^*)$:	5.47	40.21	64.98	81.07	83.51	37.39	24.39	18.08

We observe from Fig.1 that (i) the local maximum value of N^* is moving from left to right as μ increases; (ii) as λ is fixed, N^* is getting larger as μ increases. From Fig.2, we observe that (i) N^* is increasing in μ; (ii) if μ is small enough, N^* increases rapidly; (iii) if μ is large and $\frac{\lambda}{\mu}$ is small enough N^* is insensitive; and (iv)if μ is fixed and large enough, N^* increases in λ.

Case3: We select λ=0.3,μ=1.5,β=4,θ=3, α=0.4,0.8,1.2,1.6,δ=3 and vary the values of γ.

Case4: We select λ=0.5,μ=1.5,β=4,θ=3, γ =1, 2, 3, 4, δ=3 and vary the values of α.

Numerical results of case3 and case4 are presented in Table 2

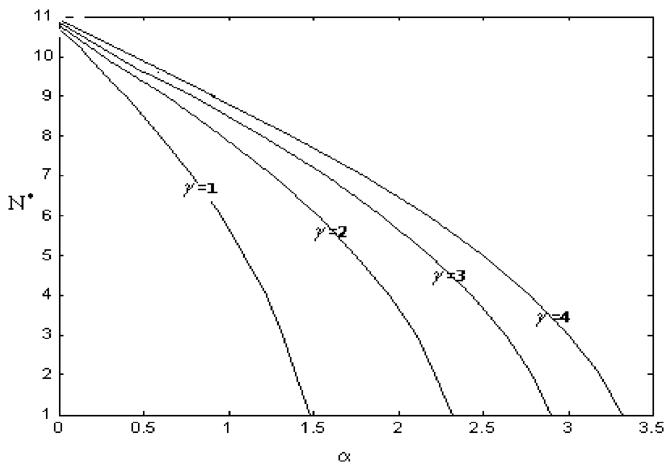

Fig.3. plots of (α, N^*) with λ=0.5, β=4, θ=3, γ =1,2,3,4, δ=3, C_h =5, C_0 =50, C_m =100, C_b =100, C_r =40, C_s =1000,.

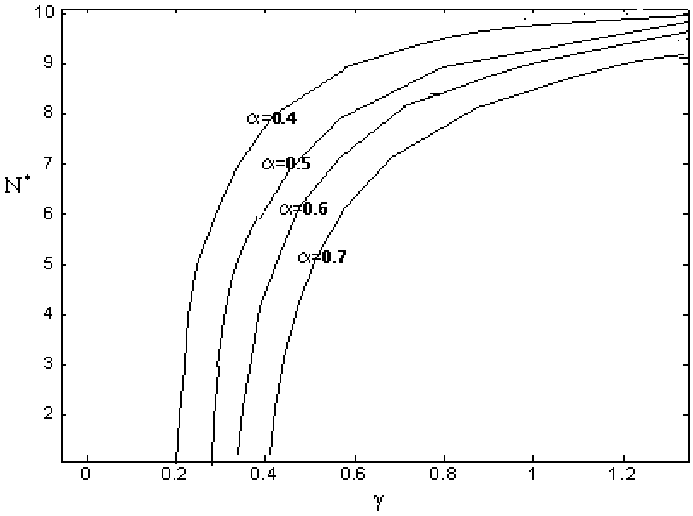

Fig.4. plots of (γ, N^*) with λ=0.5, μ=1.5, β=4, θ=3, α=0.4, 0.5, 0.6, 0.7, γ=3, δ=3, C_h=5, C_0=50, C_m=100, C_b=100, C_r=40, C_s=1000

Table2. The optimal N^* and minimum expected cost $T(N^*)$ with various (α, γ)

λ=0.5, μ=1.5, β=4, θ=3, α=0.5, δ=3, C_h=5, C_0=50, C_m=100, C_b=100, C_r=40, C_s=1000								
(α, γ)	(0.4, 3)	(0.8, 3)	(1.2, 3)	(1.6, 3)	(1.2, 1)	(1.2, 1.5)	(1.2, 2)	(1.2, 3)
N^*:	10	10	9	8	5	7	8	9
$T(N^*)$:	49.73	59.07	68.71	78.64	125.79	92.28	82.04	72.97

We observe from Fig.3 that (i) N^* decreases in α. As α is fixed, the larger γ has larger N^*; (ii) N^* has an upper bound as α closes to zero; and (iii) N^* is not insensitive to α. It can be easily observed from Fig.4 that (i) N^* increases in γ but N^* is insensitive to γ as γ is large; (ii) as γ is fixed, the larger α has smaller N^*.

Case 5: We select λ=0.5, μ=1.5, γ=4, α=0.5, δ=3 and vary the values of β and θ. Numerical results of case 5 are provided in Table 3.

Table3. The optimal N^* and minimum expected cost $T(N^*)$ with various (β, θ)

λ=0.5, μ=1.5, γ=4, α=0.5, δ=3, C_h=5, C_0=50, C_m=100, C_b=100, C_r=40, C_s=1000								
(β,θ)	(2, 3)	(3, 3)	(4, 3)	(5, 3)	(3, 1)	(3, 2)	(3, 3)	(3, 4)
N^*:	9	9	9	9	9	9	9	9
$T(N^*)$:	71.12	68.49	68.71	68.25	71.05	69.86	69.49	69.29

We observe from Table 3 that N* is insensitive with respect to changes in β and θ, where as T(N*) decreases with increase in the values β and θ.
Case 6: To observe the changes in N* and T(N*), corresponding to the changes in the cost parameters, we select λ=0.8, μ=2.5, β=3, θ=3, α=0.5, γ=3, δ=3, C_0=50, C_m=200, C_b=100, C_r=60 and vary the values of (C_s, C_h). Numerical results are presented in the Table 4. We can observe that N* decreases in C_h and increases in C_s, where as T(N*) increases in both cases.

Table4. The optimal N^* and minimum expected cost $T(N^*)$ with various (C_s, C_h)

λ=0.8, μ=2.5, β=3, θ=3, α=0.5, γ=3, δ=3, C_0=50, C_m=200, C_b=100, C_r=60								
(C_s, C_h)	(100,5)	(100,10)	(100,15)	(100,20)	(400, 10)	(600, 10)	(800,10)	(900,10)
N^*:	13	9	7	6	6	7	8	9
$T(N^*)$:	64.82	85.94	103.66	119.33	68.25	75.03	80.82	83.29

Case 7: To observe the changes in N* and T(N*), corresponding to the changes in the cost parameters C_m and C_r, we select λ=0.8, μ=2.5, β=3, θ=3, α=0.5, γ=3, δ=3, C_h=5, C_0=50, C_b=100, C_s=1000. Numerical results are presented in the Table 5. This Table reveals that N* is insensitive to C_m and C_r where as T(N*) decreases in C_r and increase in C_m.

Table5. The optimal N^* and minimum expected cost $T(N^*)$ with various (C_m, C_r)

λ=0.8, μ=2.5, β=3, θ=3, α=0.5, γ=3, δ=3, C_h=5, C_0=50, C_b=100, C_s=1000								
(C_m, C_r)	(200,30)	(200,40)	(200,50)	(200,60)	(150,60)	(170,60)	(190,60)	(210,60)
N^*:	13	13	13	13	13	13	13	13
$T(N^*)$:	73.83	70.83	67.82	64.82	64.51	64.63	64.75	64.88

10. Conclusions

Two–phase N-policy M/M/1 queueing system with server startup times, balking, breakdowns and delayed repair is studied. The explicit expressions for the steady state distribution of the number of customers in the system when the server is at different states are obtained and hence the expected system length is derived. Expected waiting time in the queue is obtained through heuristic approach. Total expected cost function for the system is formulated and determined the optimal value of the control parameter N that minimizes the expected cost. Sensitivity analysis is provided to discuss how the system performance measures will be affected by the changes in the input parameters for the investigated queueing service model.

References

[1] B.Srinivasa Kumar, V. Vasanta Kumar and T.Srinivasa Rao., (2013), Optimal control of an N-policy two-phase $M^X/M/1$ gated queueing system with server startup, subject to the server breakdowns and delayed repair, Far East Journal of Mathematical Sciences, special vol, 461-477.