Journal of Mathematical Modeling
Vol. 7, No. 2, 2019, pp. 153-174

JMM

Analysis of a queue with joining strategy and interruption repeat or resumption of service

Dhanya Shajin[*]

Department of Mathematics, Sree Narayana College, Chempazhanthy, Thiruvananthapuram, Kerala-695587, India
email: dhanya.shajin@gmail.com

Abstract. Consider an $M/M/1$ queueing system with service interruption. If the server is busy at the arrival epoch, the arriving customer decides to join the queue with probability q and balk with probability $1-q$. The service is assumed to get interrupted according to a Poisson process. The interrupted service is either resumed or restarted according to the realization of two competing independent, non-identically distributed random variables, the realization times of which follow exponential distributions. An arriving customer, finding the server under interruption does not join the system. We analyze the Nash equilibrium customers' joining strategies and give some numerical examples.

Keywords: joining strategy, interruption, repeat or resumption of service, Nash equilibrium
AMS Subject Classification: 60K25, 68M20, 90B22.

1 Introduction

In day to day life, processes like internet banking face interruptions due to power failure affecting banking procedure or the working of certain machinery. Though facilities have improved, resulting in reduced frequency of their occurrence, increasing demands (traffic) still create a jam and thereby

[*]Corresponding author.
Received: 12 October 2018 / Revised: 18 April 2019 / Accepted: 18 April 2019.
DOI: 10.22124/jmm.2019.11467.1194

interruption. In such situations, we hope to minimize its frequency by introducing some protection mechanism. Analysis of queueing models with service interruption and protection is therefore important.

Queueing model with service interruption was first studied in [22], in the context of pre-emptive priority. Inter-arrival and service times of the two type of customers are independent non-identically distributed and mutually independent random variables. The duration of interruption of a low priority customer depends on length of the busy period generated by service of the high priority customer who interrupted the service of that lower priority customer. Following this, there has been an extensive study on such models by numerous authors; we refer to the paper [11]. In [12–15] the authors studied a few queueing models with interruptions, where a special stress was given in deciding whether to restart or resume an interrupted service. Since interruption causes a natural increase in the length of a service, we cannot expect the new customers to join in the queue when the server is interrupted. This is why we included the possibility of customer loss and arriving customers decide whether to join or to balk the system.

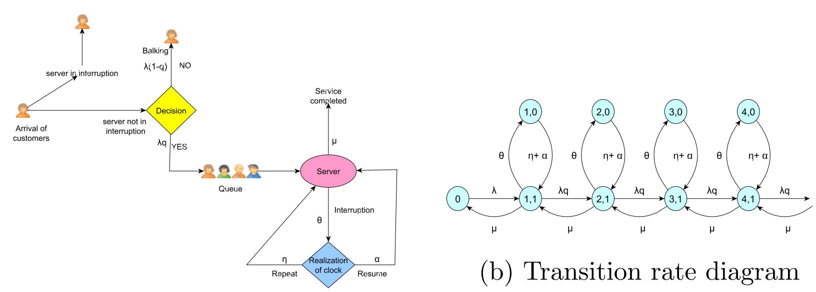
An $M/M/1$ queueing system where the customers are informed the queue length along with server's status and obtained customers' individual optimal strategy is considered in [18]. On the other hand, when customers in the system are considered as a whole, there may exist an optimal strategy. It was showed that the individually optimal strategy is different from the overall optimal strategy. This work is extended by assuming that customers cannot observe the length of queue upon their arrivals in [5]. Since then, there is a large amount of work analysing the strategic behaviour of customers such as [1–4, 6–10, 16, 17, 20, 21, 23].

As a motivating example for the present model, consider a person browsing internet for some purpose like money transactions through Net banking. The process of fund transaction may consist of a few steps, during which it may get interrupted due to different reasons. After some random time, it may happen that the interrupted service get resumed from the same step at which interruption occurred or sometimes one sees the message like "session expired" or "web page has been expired", so that the whole process is to be repeated from the beginning.

The rest of the paper is organized as follows. The model description is given in Section 2. Section 3 presents some important system characteristics. Section 4 devoted to the analysis of Nash equilibrium. Also numerical experiments are carried out in this section. Special case of the model is included in Section 5.

2 Model formulation

We consider a single server queueing system in which customers arrive according to a Poisson process with rate λ and service time is exponentially distributed with parameter μ. This service is interrupted at an exponentially distributed duration with parameter θ. At the epoch when an interruption occurs, two random clocks (resume clock and repeat clock) are started, realization times of which follow exponential distribution with parameters α and η, respectively. If the realization of the resume clock occurs first, the interrupted service is resumed whereas if the repeat clock realizes first then the interrupted service has to be repeated. Upon arrival, the arriving customer finds the service is going on, he may join the system with probability q $(0 \leq q \leq 1)$ or may leave the system with the complementary probability $1-q$. While service is stopped due to interruption, no new customer joins the system. We also assume that, when a service is interrupted, no further interruption befalls on that until the present interruption is cleared. The graphical description of this system is presented in Figure 1(a). By the assumption of the queueing system, the state of the system



(a) Pictorial representation of the model

(b) Transition rate diagram

Figure 1: Pictorial representation

under consideration can be described by a continuous time Markov chain $\Omega = \{(N(t), C(t)), t \geq 0\}$, where $N(t)$ denotes the number of customers in the system and $C(t)$ is the status of the server at time t. That is,

$$C(t) = \begin{cases} 0, & \text{interrupted service,} \\ 1, & \text{service is going on without interruption.} \end{cases}$$

The state space is $\{0\} \cup \{(n, 0), n \geq 1\} \cup \{(n, 1), n \geq 1\}$, where $\{0\}$ represents no customer in the system and the server is idle. Thus the infinitesimal generator is of the form

$$\mathcal{Q} = \begin{pmatrix} A_{00} & A_{01} & & & & \\ A_{10} & A_1 & A_0 & & & \\ & A_2 & A_1 & A_0 & & \\ & & A_2 & A_1 & A_0 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix},$$

where

$$A_{00} = [-\lambda], \quad A_{01} = \begin{bmatrix} 0 & -\lambda \end{bmatrix}, \quad A_{10} = \begin{bmatrix} 0 \\ \mu \end{bmatrix},$$

$$A_0 = \begin{bmatrix} 0 & 0 \\ 0 & \lambda q \end{bmatrix}, \; A_1 = \begin{bmatrix} -(\eta+\alpha) & (\eta+\alpha) \\ \theta & -(\lambda q+\mu+\theta) \end{bmatrix}, \; A_2 = \begin{bmatrix} 0 & 0 \\ 0 & \mu \end{bmatrix}.$$

The transition rate diagram is shown in Figure 1(b).

2.1 Stability condition

We examine the system stability. Define $A = A_0 + A_1 + A_2$. Then

$$A = \begin{bmatrix} -(\eta+\alpha) & (\eta+\alpha) \\ \theta & -\theta \end{bmatrix}.$$

This is the infinitesimal generator of the finite state continuous time Markov chain. Let $\boldsymbol{\pi} = (\pi(0), \pi(1))$ be the steady state probability vector of A. Then we have $\boldsymbol{\pi} A = \mathbf{0}, \; \boldsymbol{\pi}\mathbf{e} = 1$. This leads to

$$\pi(0) = \frac{\theta}{\theta+\eta+\alpha} \text{ and } \pi(1) = \frac{\eta+\alpha}{\theta+\eta+\alpha}.$$

The following theorem provides the stability condition of the queueing system under study.

Theorem 1. *The system under study is stable if and only if (see Theorem 3.1.1 in [19]) the left drift rate exceeds the right drift rate. That is,*

$$\lambda q < \mu. \tag{1}$$

2.2 Stationary distribution

Under the stability condition (1), we have to find the stationary distributions

$$\begin{aligned} p_0 &= \lim_{t\to\infty} P(N(t)=0), \\ p_{n0} &= \lim_{t\to\infty} P(N(t)=n, C(t)=0), \quad \text{for } n \geq 1, \\ p_{n1} &= \lim_{t\to\infty} P(N(t)=n, C(t)=1), \quad \text{for } n \geq 1. \end{aligned}$$

By the assumption of the model, we can get the steady state equation as follows:

$$\begin{aligned} \mu p_{11} &= \lambda p_0, & & \\ \theta p_{n1} &= (\eta+\alpha) p_{n0}, & n \geq 1, & \\ \lambda p_0 + (\eta+\alpha) p_{10} + \mu p_{21} &= (\lambda q + \mu + \theta) p_{11}, & & \\ \lambda q p_{n-11} + (\eta+\alpha) p_{n0} + \mu p_{n+11} &= (\lambda q + \mu + \theta) p_{n1}, & n \geq 2. & \end{aligned} \tag{2}$$

From (2) we have

$$p_{n0} = \frac{\theta}{\eta+\alpha} \left(\frac{\lambda q}{\mu}\right)^n \frac{p_0}{q}, \text{ for } n \geq 1,$$

$$p_{n1} = \left(\frac{\lambda q}{\mu}\right)^n \frac{p_0}{q}, \text{ for } n \geq 1.$$

Using the normalizing condition yields

$$p_0 = \left[\frac{(\eta+\alpha)(\mu-\lambda q)}{\delta - \lambda q(\eta+\alpha)}\right],$$

where $\delta = (\mu+\lambda)(\eta+\alpha) + \lambda\theta$.

3 Some important performance measures

- Probability that there is no customer in the system

$$p_0 = \left[\frac{(\eta+\alpha)(\mu-\lambda q)}{((\mu+\lambda)(\eta+\alpha)+\lambda\theta) - \lambda q(\eta+\alpha)}\right].$$

- Probability that there are n (≥ 1) customers in the system

$$p_n = p_{n0} + p_{n1} = \left(\frac{\lambda q}{\mu}\right)^n \left[\frac{\theta+\eta+\alpha}{q}\right] \left[\frac{(\mu-\lambda q)}{(\delta - \lambda q(\eta+\alpha))}\right].$$

- Expected number of customers in the system

$$E_N = \sum_{n=1}^{\infty} n\left(p_{n0} + p_{n1}\right) = \frac{\lambda}{\mu - \lambda q} \left[\frac{\mu(\theta+\eta+\alpha)}{(\delta - \lambda q(\eta+\alpha))}\right].$$

- Effective arrival rate $\lambda_A = \lambda \left[\dfrac{\mu(\eta+\alpha)}{(\delta - \lambda q(\eta+\alpha))}\right]$.

- Loss probability of customers $p_L = \sum_{n=1}^{\infty} p_{n0}$.
- Expected sojourn time $E_W = \dfrac{E_N}{\lambda_A} = \dfrac{\theta+\eta+\alpha}{(\eta+\alpha)(\mu-\lambda q)}$.
- Expected number of customers in the system when the server is under interruption $E_{NI} = \sum_{n=1}^{\infty} np_{n0} = \dfrac{\theta\lambda}{\mu-\lambda q}\left[\dfrac{\mu}{(\delta-\lambda q(\eta+\alpha))}\right]$.
- Expected number of customers in the system when the service is going on without interruption $E_{NS} = \sum_{n=1}^{\infty} np_{n1} = \dfrac{(\eta+\alpha)\lambda}{\mu-\lambda q}\left[\dfrac{\mu}{(\delta-\lambda q(\eta+\alpha))}\right]$.
- Variance of the number of customers in the system
$$V_N = \sum_{n=1}^{\infty} n^2 p_n - \left[\sum_{n=1}^{\infty} np_n\right]^2.$$
- Probability that the system is under interruption $P_0^* = \dfrac{\theta\lambda}{(\delta-\lambda q(\eta+\alpha))}$.
- Probability that the system is in service $P_1^* = \dfrac{(\eta+\alpha)\lambda}{(\delta-\lambda q(\eta+\alpha))}$.
- Effective interruption rate $E_{IR} = \theta P_1^*$.
- Effective rate of repetition of service $E_\eta = \eta P_0^*$.
- Effective resumption rate $E_\alpha = \alpha P_0^*$.

3.1 Expected number of interruptions during a single service

We consider the Markov chain $\{M(t), t \geq 0\}$ where $M(t)$ is the number of interruptions occurred during the service process till time t. The state space $\{0^*\} \cup \{0, 1, 2, \ldots\}$, where $\{0^*\}$denotes the absorbing state denoting the service completion. Thus the infinitesimal generator matrix is of the form

$$\mathcal{N} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ \mu & -(\mu+\theta) & \theta & & & \\ \mu & & -(\mu+\theta) & \theta & & \\ \mu & & & -(\mu+\theta) & \theta & \\ \vdots & & & & \ddots & \ddots \end{bmatrix}$$

with initial probability vector $\boldsymbol{\xi}' = (1 \;\; 0 \;\; 0 \;\; 0 \;\; \cdots)$.

Let x_n be the probability that the number of interruptions during a single service is n. Thus x_n is given by

$$x_0 = \frac{\mu}{\mu+\theta},$$

$$x_n = \frac{\mu}{\mu+\theta}\left(\frac{\theta}{\mu+\theta}\right)^n, \text{ for } n \geq 1.$$

The expected number of interruptions during any particular service is

$$E_I = \sum_{n=1}^{\infty} n x_n = \frac{\mu}{\mu+\theta}\sum_{n=1}^{\infty} n\left(\frac{\theta}{\mu+\theta}\right)^n = \frac{\theta}{\mu}.$$

Since the mean duration of an interruption is $1/(\eta+\alpha)$, hence we have the following theorem.

Theorem 2. *Expected time spent under interruption during each service is given by* $E_T^{int} = \dfrac{\theta}{(\eta+\alpha)\mu}$.

3.2 Expected service time

The service process with interruption can be viewed as a Markov chain $\{C(t), t \geq 0\}$, where $C(t)$ is the status of the server which is 0 if the service is interrupted and 1 otherwise at time t. The state space is given by $\{0,1\} \cup \{\Delta\}$, where $\{\Delta\}$ denotes the absorbing state which represents the service completion. Thus the infinitesimal generator is of the form

$$\mathcal{T} = \begin{bmatrix} B & B^0 \\ \mathbf{0} & 0 \end{bmatrix},$$

where

$$B = \begin{bmatrix} -(\eta+\alpha) & -(\eta+\alpha) \\ \theta & -(\theta+\mu) \end{bmatrix}, B^0 = \begin{bmatrix} 0 \\ \mu \end{bmatrix}.$$

Also $B\mathbf{e} + B^0 = 0$ and the initial probability vector of the process is $\boldsymbol{\xi} = (0\ 1)$. Thus the expected service time is $E_T = -\boldsymbol{\xi} B^{-1}\mathbf{e}$. Hence we have the following lemma.

Lemma 1. *Expected service time is given by* $E_T = \dfrac{\theta+\eta+\alpha}{\mu(\eta+\alpha)}$.

4 Analysis of Nash equilibrium joining strategy

In this section, we have to analyze the Nash equilibrium joining strategy under a given reward cost structure because customers have the right to decide whether to enter the system or not. The reward cost structure considered is as follows:

1. R : each customer receiving a reward of R units for completing service,
2. h : a waiting cost of h units per time unit where a customer remains in the system.

Theorem 3. *In the $M/M/1$ queueing system with service interruption repeat or resumption of service, a unique mixed strategy 'enter with probability q_e' exists and it is given by*

$$q_e = \begin{cases} 0, & \text{if} \quad R \leq \frac{h}{\mu}\left[\frac{\theta+\eta+\alpha}{\eta+\alpha}\right], \\ q_e^*, & \text{if} \quad \frac{h}{\mu}\left[\frac{\theta+\eta+\alpha}{\eta+\alpha}\right] < R < h\left[\frac{1}{\mu-\lambda} + \frac{\theta}{(\eta+\alpha)(\mu-\lambda)}\right], \\ 1, & \text{if} \quad h\left[\frac{1}{\mu-\lambda} + \frac{\theta}{(\eta+\alpha)(\mu-\lambda)}\right] \leq R, \end{cases} \tag{3}$$

where

$$q_e^* = \frac{R(\eta+\alpha)\mu - h(\theta+\eta+\alpha)}{\lambda R(\eta+\alpha)}.$$

Proof. When a tagged customer decides to enter the system at his arrival instant, the expected net benefit

$$f(q) = R - hE_W = R - h\left[\frac{\theta+\eta+\alpha}{(\eta+\alpha)(\mu-\lambda q)}\right],$$

is strictly decreasing on the interval $[0,1]$ since

$$f'(q) = -\frac{h\lambda(\theta+\eta+\alpha)}{(\eta+\alpha)(\mu-\lambda q)^2} < 0.$$

(i) If

$$f(0) = R - \frac{h}{\mu}\left[\frac{\theta+\eta+\alpha}{\eta+\alpha}\right] \leq 0,$$

the maximum benefit is non-positive which implies customers do not enter the system even if there is no customer in front of him.

(ii) If

$$\frac{h}{\mu}\left[\frac{\theta+\eta+\alpha}{\eta+\alpha}\right] < R < h\left[\frac{1}{\mu-\lambda} + \frac{\theta}{(\eta+\alpha)(\mu-\lambda)}\right],$$

then there exists a unique $q_e^* \in (0,1)$ such that $f(q) = 0$ so

$$q_e^* = \frac{R(\eta+\alpha)\mu - h(\theta+\eta+\alpha)}{\lambda R(\eta+\alpha)}.$$

This means that the tagged customer is indifferent between joining and balking the system.

(iii) If

$$f(1) = R - h\left[\frac{1}{\mu-\lambda} + \frac{\theta}{(\eta+\alpha)(\mu-\lambda)}\right] \geq 0,$$

the customers prefer to enter the system because the minimal benefit is non-negative.

Thus (i), (ii) and (iii) imply when the joining probability q adopted by other customers is smaller than q_e, the expected net benefit of an arriving customer is positive if he chooses to join the system, thus the unique best response is 1. Conversely, the unique best response is 0 if $q > q_e$ because that the expected net benefit is negative. If $q = q_e$, every strategy is the best response since the expected net benefit is always 0. This behaviour illustrates a situation that an individuals best response is a decreasing function $f(q)$ of the strategy selected by other customers. Therefore, we have to avoid the crowd situation. □

Revenue maximization

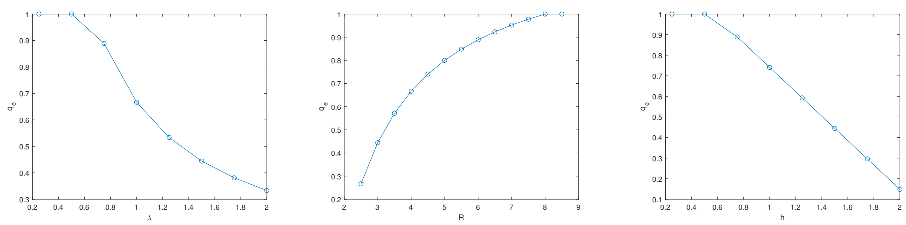
Equation (3) shows that there exists an equilibrium strategy q_e for customers, which depends on the value of p. It is readily seen that q_e is strictly decreasing in p, which means that a large proportion of customers will choose to enter the system for service if a lower p is levied. We have to find an optimal price p to maximize the revenue of the server given by $k(p) = p\lambda_A - \ell p_L - c_1 E_\eta - c_2 E_\alpha$, where c_1 is the cost due to repetition of service, c_2 is the cost due to resumption of service and ℓ is the loss cost of each lost customer.

Social optimal strategy

Now we proceed to the socially optimal strategy. For a given price p and a joining probability q, the surplus of all customers is $S_1 = \lambda_A(R-p) - hE_N$ and the server revenue is $S_2 = p\lambda_A - \ell p_L - c_1 E_\eta - c_2 E_\alpha$. Thus the expected social welfare per time unit $S(q) = S_1 + S_2$.

4.1 Numerical illustrations

In this section, we show the tendency of joining probability q_e with respect to the arrival rate λ, customer reward R and waiting cost h.



(a) q_e versus λ (b) q_e versus R (c) q_e versus h

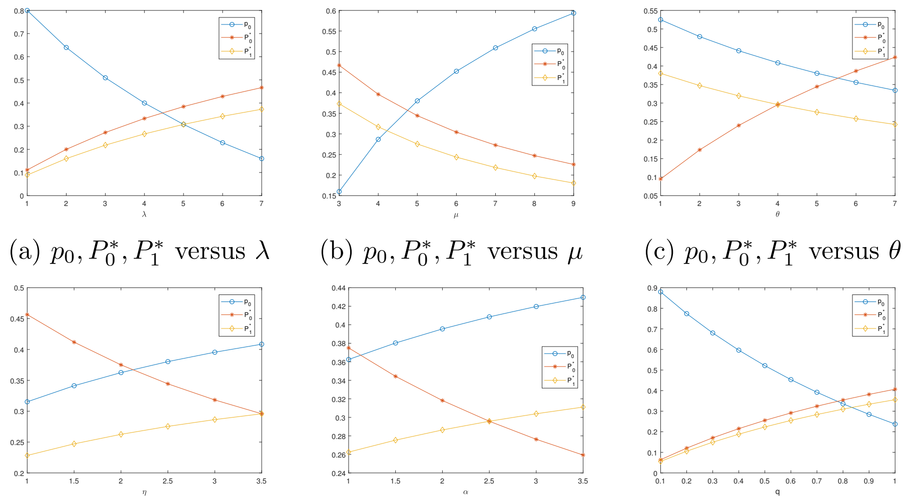
Figure 2: The joining probability q_e versus λ, R, h separately, where for Figure (a) fix $(\mu, \theta, \eta, \alpha, R, h) = (2, 2.5, 1, 0.5, 3, 1.5)$, for Figure (b) take $(\lambda, \mu, \theta, \eta, \alpha, h) = (1.5, 2, 2.5, 1, 0.5, 1.5)$, for Figure (c) put $(\lambda, \mu, \theta, \eta, \alpha, R) = (1.5, 2, 2.5, 1, 0.5, 3)$.

Figure 2 shows that for a greater value of λ, less customers are willing to join the queue, so q_e is decreasing in Figure 2(a). The same situation can be seen in Figures 2 (c). Due to the high reward R, q_e is increasing in Figure 2 (b).

From Figure 3 we see that p_0 decrease in increasing values of λ, θ, q. However, increases with increasing values of μ, η, α. Both probabilities P_0^* and P_1^* simultaneously increase in λ, q (see Figures 3(a), 3(f)) and decrease in μ, θ (see Figures 3(b), 3(c)). However, in Figures 3(d), 3(e) one of the probability increases and the other one decreases with increasing values of η and α. Effect of parameters $\lambda, \mu, \theta, \eta, \alpha, q$ on some performance measures such as expected number of customers in the system E_N, variance of the number of customers in the system V_N, expected number of customers in the system when the server is under interruption E_{NI}, expected number of customers in the system when the service is going on without interruption E_{NS} are given in Figure 4.

5 Special cases

In this section, we provide some special cases. First we consider classical queue with joining probability q, $0 \leq q \leq 1$ and no interruption in service. Secondly, an $M/Cox_2/1$ queue with interruption is considered.



(a) p_0, P_0^*, P_1^* versus λ (b) p_0, P_0^*, P_1^* versus μ (c) p_0, P_0^*, P_1^* versus θ

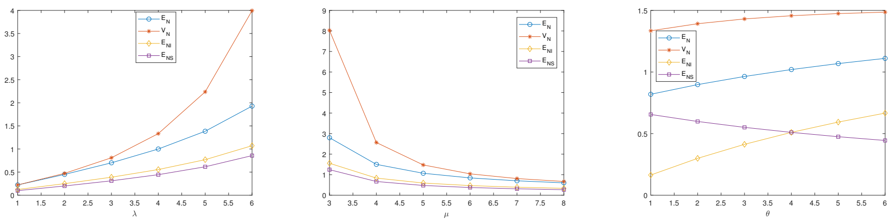
(d) p_0, P_0^*, P_1^* versus η (e) p_0, P_0^*, P_1^* versus α (f) p_0, P_0^*, P_1^* versus q

Figure 3: The probabilities p_0, P_0^*, P_1^* versus $\lambda, \mu, \theta, \eta, \alpha, q$: For Figure (a) fix $(\mu, \theta, \eta, \alpha, q) = (7, 5, 2.5, 1.5, 0.7)$, for Figure (b) take $(\lambda, \theta, \eta, \alpha, q) = (3, 5, 2.5, 1.5, 0.7)$, for Figure (c) fix $(\lambda, \mu, \eta, \alpha, q) = (3, 5, 2.5, 1.5, 0.7)$, for Figure (d) fix $(\lambda, \mu, \theta, \alpha, q) = (3, 5, 5, 1.5, 0.7)$, for Figure (e) fix $(\lambda, \mu, \theta, \eta, q) = (3, 5, 5, 1.5, 0.7)$, for Figure (f) take $(\lambda, \mu, \theta, \eta, \alpha) = (3, 5, 4, 2, 1.5)$.

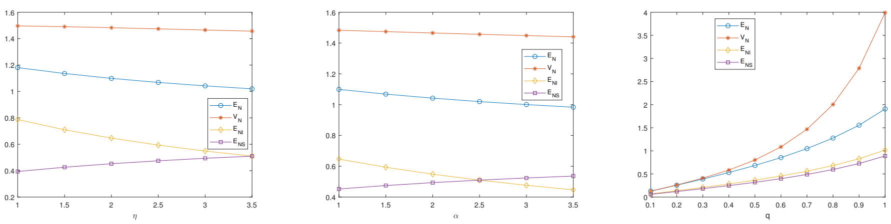
5.1 Case: 1

We consider a single server queueing system in which the service time follows exponential distribution with parameter μ. The customers arrive to the system according to a Poisson process with rate λ. If the server is busy at the arrival epoch, the arriving customer decides to join the queue with probability q and balk with probability $1-q$. The state of the system can be described by a continuous time Markov chain $\{N(t), t \geq 0\}$, where $N(t)$ denotes the number of customers in the system at time t with state space $\{0, 1, 2, \ldots\}$. Thus the infinitesimal generator is of the form

$$\mathcal{Q}' = \begin{pmatrix} -\lambda & \lambda & & & \\ \mu & -(q\lambda+\mu) & q\lambda & & \\ & \mu & -(q\lambda+\mu) & q\lambda & \\ & & \ddots & \ddots & \ddots \end{pmatrix}.$$



(a) E_N, V_N, E_{NI}, E_{NS} versus λ (b) E_N, V_N, E_{NI}, E_{NS} versus μ (c) E_N, V_N, E_{NI}, E_{NS} versus θ



(d) E_N, V_N, E_{NI}, E_{NS} versus η (e) E_N, V_N, E_{NI}, E_{NS} versus α (f) E_N, V_N, E_{NI}, E_{NS} versus q

Figure 4: Effect of E_N, V_N, E_{NI}, E_{NS} versus $\lambda, \mu, \theta, \eta, \alpha, q$: For Figure (a) fix $(\mu, \theta, \eta, \alpha, q) = (7, 5, 2.5, 1.5, 0.7)$, for Figure (b) take $(\lambda, \theta, \eta, \alpha, q) = (3, 5, 2.5, 1.5, 0.7)$, for Figure (c) fix $(\lambda, \mu, \eta, \alpha, q) = (3, 5, 2.5, 1.5, 0.7)$, for Figure (d) fix $(\lambda, \mu, \theta, \alpha, q) = (3, 5, 5, 1.5, 0.7)$, for Figure (e) fix $(\lambda, \mu, \theta, \eta, q) = (3, 5, 5, 1.5, 0.7)$, for Figure (f) take $(\lambda, \mu, \theta, \eta, \alpha) = (3, 5, 4, 2, 1.5)$.

The system under study is stable if and only if $\lambda q < \mu$. Under the stability condition $\lambda q < \mu$, we have to find the stationary distributions

$$\xi_n = \frac{\lambda}{\mu}\left(\frac{q\lambda}{\mu}\right)^{n-1}\xi_0 \quad \text{with} \quad \xi_0 = \frac{\mu - q\lambda}{\mu + (1-q)\lambda}.$$

- Probability that there is no customer in the system, $\xi_0 = \dfrac{\mu - \lambda q}{\mu + (1-q)\lambda}$.
- Probability that there are n (≥ 1) customers in the system,

$$\xi_n = \frac{\lambda}{\mu}\left(\frac{q\lambda}{\mu}\right)^{n-1}\left[\frac{\mu - \lambda q}{\mu + (1-q)\lambda}\right].$$

- Expected number of customers in the system,

$$E'_N = \frac{\lambda}{\mu - q\lambda}\left[\frac{\mu}{\mu + (1-q)\lambda}\right].$$

- Effective arrival rate, $\lambda'_A = \lambda \left[\dfrac{\mu}{\mu + (1-q)\lambda}\right]$.
- Expected sojourn time, $E'_W = \dfrac{1}{\mu - \lambda q}$.

5.1.1 Analysis of Nash equilibrium joining strategy

In this section, we have to analyze the Nash equilibrium joining strategy of $M/M/1$ queueing system without interruption of service under a given reward cost structure because customers have the right to decide whether to enter the system or not. When a tagged customer decides to enter the system at his arrival instant, the expected net benefit is

$$g(q) = R - hE'_W = R - h\left[\frac{1}{\mu - q\lambda}\right],$$

where R and h are given in Section 4. $g(q)$ is strictly decreasing on the interval $[0,1]$ since

$$g'(q) = -\frac{h\lambda}{(\mu - q\lambda)^2} < 0.$$

(i) If

$$g(0) = R - \frac{h}{\mu} \leq 0,$$

the maximum benefit is non-positive which implies customers do not enter the system even if there is no customer in front of him.

(ii) If

$$\frac{h}{\mu} < R < \frac{h}{\mu - \lambda},$$

then there exists a unique $q_e^* \in (0,1)$ such that $g(q) = 0$ so

$$q_e^* = \frac{R\mu - h}{R\lambda}.$$

This means that the tagged customer is indifferent between joining and balking the system.

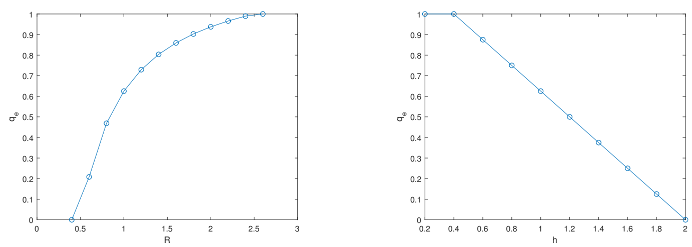
(iii) If

$$g(1) = R - \frac{h}{\mu - \lambda} \geq 0,$$

then the customers prefer to enter the system because the minimal benefit is non-negative.

Theorem 4. *In the $M/M/1$ queueing system without interruption of service, a unique mixed strategy 'enter with probability q_e' exists and it is given by*

$$q_e = \begin{cases} 0, & \textit{if} \quad R \leq \frac{h}{\mu}, \\ \dfrac{R\mu - h}{R\lambda}, & \textit{if} \quad \frac{h}{\mu} < R < \frac{h}{\mu-\lambda}, \\ 1, & \textit{if} \quad \frac{h}{\mu-\lambda} \leq R. \end{cases}$$



(a) q_e versus R (b) q_e versus h

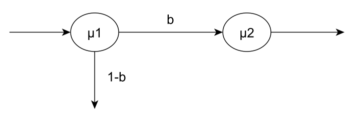
Figure 5: The joining probability q_e versus R and h separately where for Figure (a) take $(\lambda, \mu, h) = (0.8, 1, 0.5)$, for Figure (b) put $(\lambda, \mu, R) = (0.8, 1, 2)$.

Figure 5 shows that for a greater value of h, q_e is decreasing in Figure 5(b). Due to the high reward R, we obtain q_e is increasing in R from the Figure 5(a).

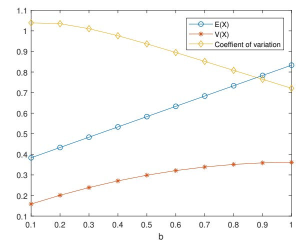
5.2 Case: 2

In this section, we consider the service time of a customer has a coxian-2 distribution with parameter (b, μ_1, μ_2). Without loss of generality we assume that $\mu_1 \geq \mu_2$. The service mechanism may be considered as follows. The customer first goes through phase 1 to get his service completed with probability $1-b$, or goes through a second phase with probability b (see Figure 6). The sojourn time in the two phases are independent exponential random variable with mean $1/\mu_1$ and $1/\mu_2$, respectively. The expectation $(E[X])$, variance $(V[X])$ and coefficient of variation (cov) of coxian-2 distribution is $\frac{1}{\mu_1} + \frac{b}{\mu_2}, \frac{1}{\mu_1^2} + \frac{b(2-b)}{\mu_2^2}, \frac{\sqrt{V[X]}}{E[X]}$, respectively.

Customers arrive according to a Poisson process with rate λ. Upon arrival, the arriving customer finds the service is going on, he may join the system with probability q $(0 \leq q \leq 1)$ or may leave the system with the complementary probability $1-q$. While the service is stopped due



(a) Phase change



(b) $E[X], V[X]$ and cov

Figure 6: Coxian-2

to interruption, no new customers join the system. The interruption occur according to a Poisson process with rate θ. The interrupted service is taken for repair immediately with repair distribution following an exponential distribution with parameter η. A random clock is started at the beginning of each repair to decide whether to restart or resume the service after repair. If the random clock realizes before a repair, the service needs to be restarted, otherwise the service is resumed in the phase from where interruption occurred. The realization time of the random clock also follows an exponential distribution with parameter γ (see Figure 7).

An example related to this model discussed can be described as follows: In general, antibiotics are prescribed for a specified duration of time. Interruptions of short durations are permitted. However, if the medicine is not taken continuously for a few days, the whole process has to be repeated.

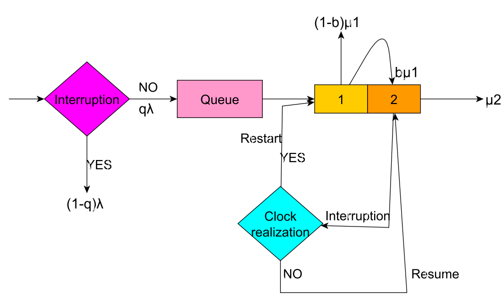


Figure 7: Pictorial representation of the model

Under the above assumption of the queueing system, the state of the system can be described by a continuous time Markov chain $\{(N(t), C(t), S(t)), t \geq 0\}$, where $N(t)$ denotes the number of customers in the system and $C(t)$

is the status of the server at time t. That is,

$$C(t) = \begin{cases} 0, & \text{service is going on without interruption} \\ 1, & \text{interrupted with a running clock} \\ 2, & \text{interrupted with a realized random clock} \end{cases}$$

$S(t)$ is the phase of the service process at time t. The state space is $\{0\} \cup \{(n,i,j), n \geq 1; i = 0,1,2; j = 1,2\}$, where $\{0\}$ represents no customer in the system and the server is idle. Thus the infinitesimal generator is of the form

$$\tilde{\mathcal{Q}} = \begin{pmatrix} \tilde{A}_{00} & \tilde{A}_{01} & & & \\ \tilde{A}_{10} & \tilde{A}_1 & \tilde{A}_0 & & \\ & \tilde{A}_2 & \tilde{A}_1 & \tilde{A}_0 & \\ & & \ddots & \ddots & \ddots \end{pmatrix},$$

where $\tilde{A}_{00} = [-\lambda]$, $\tilde{A}_{01} = \begin{bmatrix} \lambda\boldsymbol{\alpha} & \mathbf{0} & \mathbf{0} \end{bmatrix}$,

$$\tilde{A}_{10} = \begin{bmatrix} \mathbf{S}^0 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad \tilde{A}_0 = \begin{bmatrix} \lambda q I & O & O \\ O & O & O \\ O & O & O \end{bmatrix},$$

$$\tilde{A}_1 = \begin{bmatrix} S - (\lambda q + \theta)I & \theta I & O \\ \eta I & -(\eta+\gamma)I & \gamma I \\ \eta \mathbf{e}\boldsymbol{\alpha} & O & -\eta I \end{bmatrix}, \quad \tilde{A}_2 = \begin{bmatrix} \mathbf{S}^0\boldsymbol{\alpha} & O & O \\ O & O & O \\ O & O & O \end{bmatrix},$$

with

$$S = \begin{bmatrix} -\mu_1 & b\mu_1 \\ 0 & -\mu_2 \end{bmatrix}, \quad \mathbf{S}^0 = \begin{bmatrix} (1-b)\mu_1 \\ \mu_2 \end{bmatrix}, \quad \boldsymbol{\alpha} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

and $S\mathbf{e} + \mathbf{S}^0 = \mathbf{0}$.

5.2.1 Stability condition

Next we examine the system stability. Define $\tilde{A} = \tilde{A}_0 + \tilde{A}_1 + \tilde{A}_2$. Then

$$\tilde{A} = \begin{bmatrix} S + \mathbf{S}^0\boldsymbol{\alpha} - \theta I & \theta I & O \\ \eta I & -(\eta+\gamma)I & \gamma I \\ \eta \mathbf{e}\boldsymbol{\alpha} & O & -\eta I \end{bmatrix}.$$

This is the infinitesimal generator of the finite state continuous time Markov chain. Let $\tilde{\boldsymbol{\pi}} = (\tilde{\boldsymbol{\pi}}_0, \tilde{\boldsymbol{\pi}}_1, \tilde{\boldsymbol{\pi}}_2)$ be the steady state probability vector of $\tilde{A}$.

Then we have $\tilde{\boldsymbol{\pi}}\tilde{A} = \mathbf{0}$, $\tilde{\boldsymbol{\pi}}\mathbf{e} = 1$. This leads to

$$\tilde{\boldsymbol{\pi}}_i(j) = \begin{cases} \left(\dfrac{\eta}{\eta+\theta}\right)\left[\dfrac{\mu_2(\eta+\gamma)+\gamma\theta}{(b\mu_1+\mu_2)(\eta+\gamma)+\gamma\theta}\right], & i=0, \quad j=1, \\ \dfrac{b\mu_1(\eta+\gamma)}{\mu_2(\eta+\gamma)+\gamma\theta}\tilde{\boldsymbol{\pi}}_0(1), & i=0, \quad j=2, \\ \dfrac{\theta}{\eta+\gamma}\tilde{\boldsymbol{\pi}}_0(1), & i=1, \quad j=1, \\ \dfrac{\theta}{\eta+\gamma}\tilde{\boldsymbol{\pi}}_0(2), & i=1, \quad j=2, \\ \dfrac{\gamma\theta}{\eta(\eta+\gamma)}\tilde{\boldsymbol{\pi}}_0(1), & i=2, \quad j=1, \\ \dfrac{\gamma\theta}{\eta(\eta+\gamma)}\tilde{\boldsymbol{\pi}}_0(2), & i=2, \quad j=2. \end{cases}$$

The following theorem provides the stability condition of the queueing system under study.

Theorem 5. *The system under study is stable if and only if (from the Theorem 3.1.1 in Neuts [19]) the left drift rate exceeds the right drift rate. That is,*

$$\lambda q < \frac{\mu_1\mu_2(\eta+\gamma)+(1-b)\mu_1\gamma\theta}{(b\mu_1+\mu_2)(\eta+\gamma)}.$$

5.2.2 Steady-state probability vector

The stationary probability vector $\tilde{\mathbf{x}}$ is given by $\tilde{\mathbf{x}}\tilde{\mathcal{Q}} = \mathbf{0}$, $\tilde{\mathbf{x}}\mathbf{e} = 1$.

On partitioning the steady state vector as $\tilde{\mathbf{x}} = (\tilde{\mathbf{x}}_0, \tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots)$, the equation $\tilde{\mathbf{x}}\tilde{\mathcal{Q}} = \mathbf{0}$ reduces to the following equations:

$$\begin{aligned} \tilde{\mathbf{x}}_0\tilde{A}_{00} + \tilde{\mathbf{x}}_1\tilde{A}_{10} &= \mathbf{0}, \\ \tilde{\mathbf{x}}_0\tilde{A}_{01} + \tilde{\mathbf{x}}_1\tilde{A}_1 + \tilde{\mathbf{x}}_2\tilde{A}_2 &= \mathbf{0}, \\ \tilde{\mathbf{x}}_{n-1}\tilde{A}_0 + \tilde{\mathbf{x}}_n\tilde{A}_1 + \tilde{\mathbf{x}}_{n+1}\tilde{A}_2 &= \mathbf{0}, \quad n \geq 2. \end{aligned}$$

Also $\lambda\tilde{\mathbf{x}}_0 = \tilde{\mathbf{x}}_1 \begin{pmatrix} \mathbf{S}^0 \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}$ and $\lambda q\tilde{\mathbf{x}}_n\mathbf{e}' = \tilde{\mathbf{x}}_{n+1} \begin{pmatrix} \mathbf{S}^0 \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}$, $n \geq 1$.

By Theorem 3.2.1 of Neuts [19], we get $\tilde{\mathbf{x}}_0 = (1-\rho)$ and $\tilde{\mathbf{x}}_n = q(1-\rho)\boldsymbol{\beta}R^n$ for $n \geq 1$, where

$$\rho = \frac{\lambda q(b\mu_1+\mu_2)(\eta+\gamma)}{\mu_1\mu_2(\eta+\gamma)+(1-b)\mu_1\gamma\theta},$$

$$\mathbf{e}' = \begin{pmatrix} \mathbf{e} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\alpha} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad R = -\lambda\left(\tilde{A}_1 + \lambda q\mathbf{e}'\boldsymbol{\beta}\right)^{-1}.$$

5.2.3 Expected waiting time

For computing expected waiting time of a particular customer who joins as the r^{th} customer in the queue, we consider the Markov chain

$$\{(N'(t), C(t), S(t)), t \geq 0\},$$

where $N'(t)$ is the rank of the customer,

$$C(t) = \begin{cases} 0, & \text{service is going on without interruption,} \\ 1, & \text{interrupted with a running clock,} \\ 2, & \text{interrupted with a realized random clock,} \end{cases}$$

and $S(t)$ is the phase of the service process at time t. The rank $N(t)$ of the customer is assumed to be r if he joins as the r^{th} customer in the queue. His rank decreases by 1 as one customer ahead of him leaves the system after completing the service. Since the customers who arrive after the tagged customer cannot change his rank. Thus the state space is of the form $\{(n, i, j), 0 \leq n \leq r; i = 0, 1, 2; j = 1, 2\} \cup \{\Delta\}$, where $\{\Delta\}$ is an absorbing state in the sense that the tagged customer is selected for service. The infinitesimal generator is of the form

$$\mathcal{W} = \begin{pmatrix} W & \mathbf{W}^0 \\ \mathbf{0} & 0 \end{pmatrix},$$

where

$$W = \begin{pmatrix} H_1 & A_2 & & \\ & \ddots & \ddots & \\ & & H_1 & A_2 \\ & & & H_1 \end{pmatrix}, \quad \mathbf{W}^0 = \begin{pmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ H \end{pmatrix},$$

with

$$H_1 = \begin{pmatrix} S - \theta I & \theta I & O \\ \eta I & -(\eta + \gamma) I & \gamma I \\ \eta \mathbf{e} \boldsymbol{\alpha} & O & -\eta I \end{pmatrix}, \quad H = \begin{pmatrix} \mathbf{S}^0 \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

Let τ be the time until absorption of the service process. Hence $E(\tau)$ is the expected service time until absorption, $E(\tau) = \boldsymbol{\beta} \left(-H_1\right)^{-1} \mathbf{e}$.

Now the waiting time W_T of a customer, who joins the queue in the r^{th} customer is the time until absorption in the above Markov chain. The expected waiting time of this particular customer is given by the column vector

$$E_W^r = -W^{-1}\mathbf{e} = -\left(H_1\right)^{-1}\mathbf{e} + (r-1)E(\tau)\mathbf{e}.$$

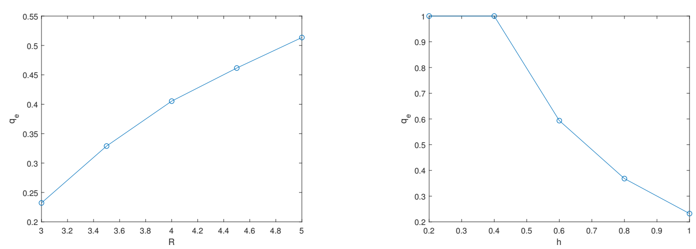
Hence, the expected waiting time of a general customer is $E_W = \sum_{r=1}^{\infty} \tilde{\mathbf{x}}_r E_W^r$.

5.2.4 Analysis of joining strategy

In this section we have to analyze the expected net benefit $g(q) = R - hE_W$, when a tagged customer decides to enter the system at his arrival instant.

5.2.5 Numerical illustrations

In this section we show the tendency of joining probability q_e with respect to the customer reward R and waiting cost h.



(a) q_e versus R (b) q_e versus h

Figure 8: The joining probability q_e versus R and h separately where for Figure (a) take $(\lambda, b, \mu_1, \mu_2, \theta, \eta, \gamma, h) = (1.5, 0.25, 2, 1, 2, 1, 0.5, 1)$, for Figure (b) put $(\lambda, b, \mu_1, \mu_2, \theta, \eta, \gamma, R) = (1.5, 0.25, 2, 1, 2, 1, 0.5, 3)$.

Figure 8 shows that for a greater value of h, q_e is decreasing in Figure 8(b). Due to the high reward R, we obtain q_e is increasing in R from the Figure 8(a).

6 Conclusions

In this paper, we studied a single-server queueing system in which the service process may face interruptions during service. The interrupted service is either resumed or restarted according to the realization of two competing independent, non-identically distributed random variables, the realization times of which follow exponential distributions. Customers arrive to the system according to a Poisson process and they are not permitted to join the system if the system is under interruption. If the server is busy at the arrival epoch, the arriving customer decides to join the queue with probability q and balk with probability $1-q$. The service times are exponentially distributed. We analyzed the Nash equilibrium customers' joining strategies and some numerical examples. This system could be extended to the case of multi-server and queueing-inventory system.

Acknowledgements

The author would like to thank the reviewers for their valuable comments and constructive suggestions that helped to improve the presentation of this paper.

References

[1] O. Boudali and A. Economou, *Optimal and equilibrium balking strategies in the single server Markovian queue with catastrophes*, European J. Oper. Res. **218** (2012) 708–715.

[2] A. Economou and S. Kanta, *Equilibrium balking strategies in the observable single-server queue with breakdowns and repairs*, Oper. Res. Lett. **36** (2008) 696–699.

[3] A. Economou, S. Kanta, *Equilibrium customer strategies and social profit maximization in the single server constant retrial queue*, Naval Res. Logist. **58** (2011) 107–122.

[4] A. Economou and A. Manou, *Equilibrium balking strategies for a clearing queueing system in alternating environment*, Ann. Oper. Res. **208** (2013) 489-514.

[5] N.M. Edelson and D. K. Hilderbrand, *Congestion tolls for Poisson queuing processes*, Econometrica **43** (1975) 81–92.

[6] A. Elcan, *Optimal customer return rate for an $M/M/1$ queueing system with retrials*, Probab. Engrg. Inform. Sci. **8** (1994) 521–539.

[7] P. Guo and R. Hassin, *Strategic behavior and social optimization in Markovian vacation queues*, Oper. Res. **59** (2011) 986–997.

[8] P. Guo and R. Hassin, *Strategic behavior and social optimization in Markovian vacation queues: The Case of Heterogeneous Customers*, European J. Oper. Res. **222** (2012) 278–286.

[9] P. Guo and Q. Li, *Strategic behavior and social optimization in partially-observable Markovian vacation queues*, Oper. Res. Lett. **41** (2013) 277–284.

[10] R. Hassin and M. Haviv, *On optimal and equilibrium retrial rates in a queueing system*, Probab. Engrg. Inform. Sci. **10** (1996) 223–227.

[11] A. Krishnamoorthy and P.K. Pramod, *Queues with interruption and repeat/resumption of service-a survey and further results*, fourth international conference on neural, Parallel & Scientific Computations; August 11–14, Morehouse College, Atlanta, Georgia, USA, 2010.

[12] A. Krishnamoorthy, P.K. Pramod and T.G. Deepak, *On a queue with interruptions and repeat or resumption of service*, Nonlinear Anal. Theory Methods Appl. **71** (2009) 1673–1683.

[13] A. Krishnamoorthy, B. Gopakumar and N.C. Viswanath, *A queueing model with interruption resumption/restart and reneging*, Bulletin of Kerala Mathematics Association, Special Issue, 29–45, FRS, 2009.

[14] A. Krishnamoorthy, B. Gopakumar and N.C. Viswanath, *An $M/E_m/1$ queue with protected and unprotected phases from interruption*; presented at 5^{th} international conference on queueing theory and network applications; July 24–26, Beijing, China, 2010.

[15] A. Krishnamoorthy, S.S. Nair and N.C. Viswanath, *An inventory model with server interruptions*, presented at 5^{th} international conference on queueing theory and network applications; July 24–26, Beijing, China, 2010.

[16] X.Y. Li, J.T. Wang and F. Zhang, *New results on equilibrium balking strategies in the single-server queue with breakdowns and repairs*, Appl. Math. Comput. **241** (2014) 380–388.

[17] W.Q. Liu, Y. Ma and J.H. Li, *Equilibrium threshold strategies in observable queueing systems under single vacation policy*, Appl. Math. Model. **36** (2012) 6186–6202.

[18] P. Naor, *The regulation of queue size by levying tolls*, Econometrica **37** (1969) 15–24.

[19] M.F. Neuts, *Matrix-geometric solutions in stochastic models an algorithmic approach*, Johns Hopkins University Press, Baltimore, 1981.

[20] J.T. Wang and F. Zhang, *Equilibrium analysis of the observable queues with balking and delayed repairs*, Appl. Math. Comput. **218** (2011) 2716–2729.

[21] J.T. Wang and F. Zhang, *Strategic joining in $M/M/1$ retrial queues*, European J. Oper. Res. **230** (2013) 76–87.

[22] H. White, L. Christie, *Queuing with preemptive priorities or with breakdown*, Oper. Res. **6** (1958) 79–95.

[23] F. Zhang and J.T. Wang, *On the optimal and equilibrium retrial rates in an unreliable retrial queue with vacations*, Ind. Manag. Optim. **8** (2012) 861–875.