Equilibrium Mixed Strategies in a Discrete-Time Markovian Queue Under Multiple and Single Vacation Policies

Zaiming Liu¹, Yan Ma¹ and Zhe George Zhang²,3,*

¹School of Mathematics and Statistics, Central South University, Changsha, China
²College of Business and Economics, Western Washington University, Bellingham, WA, USA
³Beedie School of Business, Simon Fraser University, British Columbia, Canada

(Received November 2013, accepted May 2014)

Abstract: We study the Nash equilibrium behavior of customers in a discrete-time single-server queue under multiple vacation policy or single vacation policy. Every arriving customer either joins the queue or balks based on his or her service utility. Using a Quasi-Birth-Death Processes (QBD) model, we obtain the stationary distribution of the queue length via the Matrix-Geometric Solution method. We analyze the equilibrium mixed threshold strategies under two different vacation policies. By presenting numerical examples, we examine the impacts of system parameters on the equilibrium customer behavior and compare the single and multiple vacation policies in terms of the social welfare.

Keywords: Economics of queues, Matrix-Geometric solution, multiple vacations, Nash equilibrium strategy, single vacation.

1. Introduction

In a classical queueing model, customers usually do not have the options of joining the queue or balking. This assumption limits the application of the queueing model. There have always been considerable interests of researchers in the analysis on queueing system with customer choice from an economic viewpoint. In such a setting, a reward-cost structure is assumed that incorporates the customers’ desire for service and their aversion to waiting. We consider the case where customers can make decisions about joining or balking. The resulting queueing model can be treated as a symmetric game among the customers, where the basic problem is to find the equilibrium strategy.

The literature on a $M/M/1$ type queueing system with strategic customers begins with Naor [14], who studies equilibrium and social optimal join-or-balk strategies under a simple linear reward-cost structure. Naor's model is further refined and extended by several authors, e.g., Yechiali [19], Johansen and Stidham Jr. [9], Stidham Jr. [17] and Mendelson and Whang [13]. Schroeter [16] considers heterogeneous customers with uniformly distributed costs. Chen and Frank [2] generalize Naor's model by assuming that both customers and server maximize their expected discounted utility using a common discount rate. Larsen [10] considers another generalization of Naor's model by assuming the customers differ by their service values. Erlichman and Hassin [4] discuss a single server Markovian queue allowing customers to overtake others. Guo et al. [7] study customer equilibrium as well as socially optimal strategies to join a queue with only partial

* Corresponding author. E-mail: gzhang@sfu.ca
There are some recent studies on the equilibrium customer behaviors in vacation queues. It is first presented by Burnetas and Economou [1], who study a system with an exponential setup time for each busy period. They consider customers’ strategic behavior under different levels of information which may include the queue length and/or the state of the server. Subsequently, Economou and Kanta [3] consider the Markovian queue that alternates between on and off periods in observable queue case. Sun et al. [18] present the equilibrium customer behavior in an observable \( M / M / 1 \) queue under interruptible and insusceptible setup/closedown policies. Guo and Hassin [5] examine an \( M / M / 1 \) system with an N-policy in which the server is triggered to work only after the queue length reaches threshold N. Recently, Guo and Zhang [6] study strategic customer behaviors in a multiserver stochastic service system with a congestion-based staffing (CBS) policy suggested by Zhang [20] under no and partial information scenarios.

In this paper, we study the mixed threshold strategies which are more general and flexible in equilibrium analysis. Pure threshold strategy is typical in the literature. For example, Liu et al.[11] and Ma et al.[12] analyze the vacation queues with the pure threshold strategies. However, while the pure threshold policy may not exist, the mixed strategy exist in equilibrium analysis. For example, the following situation is possible when customers’ net benefit is close to zero. If other customers in the population use the threshold \( Q_e \), then the best response for a tagged individual is \( Q_e + 1 \) instead of \( Q_e \) and if other customers use the threshold \( Q_e - 1 \) then this tagged customer’s best response becomes \( Q_e \) instead of \( Q_e - 1 \) (see [12]). Hence in this case, a pure strategy cannot achieve an equilibrium for all customers. However, we can identify a mixed threshold strategy \( (Q_e, q), q \in [0, 1] \) to achieve the equilibrium. It is a mixture of two pure threshold strategies \( Q_e - 1 \) and \( Q_e \) in the sense that strategy \( Q_e - 1 \) is followed with probability of 1–\( q \) and strategy \( Q_e \) is followed with probability of \( q \). Clearly, pure strategy is the special case of the mixed strategy with \( q = 0 \). We explore the mixed strategies in discrete-time vacation queueing systems and compare the social benefits in single and multiple vacation cases in this paper.

The reminder of this paper is organized as follows. Section 2 presents the description of the model. Section 3 analyzes the equilibrium behavior via Matrix Analytical Method. In Section 4, we illustrate the effects of system parameters and different vacation policies on the equilibrium states and social benefits via analytical and numerical comparisons. Finally, Section 5 concludes this paper.

2. Model Description

We consider a discrete-time single server queue with infinite capacity. Customers are only informed about the queue length information upon arrival without knowledge about the state of the server. This case is also known as the observable queue situation. In a multiple vacation queueing system, the server takes a vacation immediately at the end of each busy period. If the server finds an empty system at the end of the vacation, he will take another vacation. He keeps taking vacations until some waiting customers exist at the end of a vacation and then he resumes serving the queue. In contrast, in a single vacation queue, the server takes exactly one vacation immediately at the end of each busy period. If the server finds no customers in the system upon returning from the vacation, he becomes idle and waits for the next arriving customer. When a customer arrives, it immediately
starts serving the customer. In this paper, for any real number \( x \in [0,1] \), we denote \( \overline{x} = 1 - x \).

Assume that customer arrivals occur at the end of slot \( t = n^+ \), \( n = 0,1, \ldots \). The inter-arrival times are independent and identically distributed (i.i.d.) discrete random variables, denoted by \( T \), following a geometric distribution with rate \( p \).

\[
P(T = k) = p\overline{p}^{k-1}, \quad k \geq 1, \quad 0 < p < 1.
\]

The beginning and ending of service occur at slot division point \( t = n \), \( n = 0,1, \ldots \). The service times are i.i.d. random variables, denoted by \( S \), following a geometric distribution with rate \( \mu \).

\[
P(S = k) = \mu\overline{\mu}^{k-1}, \quad k \geq 1, \quad 0 < \mu < 1.
\]

The vacation starting and ending of vacation also occur at \( t = n^- \). The vacation times are i.i.d. random variable, denoted by \( V \), following a geometric distribution with rate \( \theta \).

\[
P(V = k) = \theta\overline{\theta}^{k-1}, \quad k \geq 1, \quad 0 < \theta < 1.
\]

We assume that inter-arrival times, service times, and vacation times are mutually independent. The service discipline is the First-Come-First-Served (FCFS). Moreover, suppose \( p < \mu \) to guarantee that the system reaches the steady state.

Let \( Q_n \) be the number of customers in the system at time \( n^+ \). According to the assumptions above, a customer with service completed leaving at \( t = n^+ \) is not counted for \( Q_n \) but a customer arriving at \( t = n^- \) is counted for \( Q_n \). We assume

\[
I_n = \begin{cases} 
0, & \text{the system is in a vacation period at time } n^+ \\
1, & \text{the system is in a service period at time } n^+.
\end{cases}
\]

It is clear that \( \{ Q_n, I_n \} \) is a two-dimensional Markov chain.

We are interested in the customer's choice behavior (joining or balking) at his arrival instant. To model the decision process, we assume that every customer receives a reward of \( R \) units for getting the service. Meanwhile, there is a waiting cost of \( C \) units per time unit for a customer in the system (in queue or in service). Customers are risk neutral and want to maximize their expected net reward (benefit). From now on, we assume

\[
R > \frac{C}{\mu} + \frac{C}{\theta}.
\]

This condition ensures that the reward for service exceeds the expected cost for a customer joining an empty system. Such a condition avoids a trivial and non-realistic situation of nobody joining the system. Finally, we assume that there are no retrials of balking customers or reneging of waiting customers.

Suppose \( B \) be the expected utility/net benefit and \( \omega \) be the mean sojourn time for a customer joining the system. Given \( k \) customers are in system, since the state of server is unobservable, there is a probability that customer's arrival occurs in the vacation time. Let \( \pi_{t|Q}(0|k) \) be the probability that a customer arrives during a server's vacation period given \( k \) customers are in system. Therefore, the mean sojourn time \( \omega = (k+1/\mu) + (\pi_{t|Q}(0|k)/\theta) \). Hence
\[
B = R - \frac{C(k+1)}{\mu} - \frac{C\pi_{i|Q}(0|k)}{\theta}.
\]

The customer prefers to joining the queue if \( B \) is positive and is indifferent between joining and balking if \( B \) equals zero. Thus it follows from \( B \geq 0 \) that there exists a certain threshold value \( Q \) such that customers would not enter when the queue length exceeds \( Q \). We also assume \( Q = Q_c + 1 \).

3. Equilibrium Analysis

In this section, we first consider the multiple vacation system, a relative simple model without the idle state; and then extend our analysis to the single vacation queueing system. Finally, we discuss the equilibrium strategies.

3.1. Multiple Vacations Policy Model

It is clear that the state space \( \Omega = \{(k,i)\mid k \geq i, i = 0, 1\} \). Using the lexicographical sequence for the states, the transition probability matrix can be written as

\[
P = \begin{bmatrix}
A & C \\
B & A_1 & C_1 \\
& B_1 & A_1 & C_1 \\
& & \ddots & \ddots & \ddots \\
& & B_1 & A_1 & C_1 \\
& & & B_1 & A_2 & C_2 \\
& & & & B_2 & A_3_{(Q_c+1,Q_c+1)} \\
\end{bmatrix},
\]

where

\[
A = A_{m} = (\bar{p}), \quad C = C_{m} = (\bar{\theta}/p, \theta/p),
\]

\[
B = B_{m} = \begin{bmatrix} 0 \\ \bar{p}/\mu \end{bmatrix}, \quad A_1 = \begin{bmatrix} \bar{\theta}/\bar{p} & \theta/\bar{p} \\ 0 & \bar{p}/(1 - \bar{p}/\mu - \bar{p}/\mu) \end{bmatrix} , \quad C_1 = \begin{bmatrix} \bar{\theta}/p & \theta/p \\ 0 & p/\mu \end{bmatrix},
\]

\[
B_1 = \begin{bmatrix} 0 & 0 \\ \bar{p}/\mu \end{bmatrix}, \quad A_2 = \begin{bmatrix} \bar{\theta}/pq & \theta/pq \\ 0 & \bar{p}/(1 - \bar{p}/\mu - pq/\mu) \end{bmatrix}, \quad C_2 = \begin{bmatrix} \bar{\theta}/pq & \theta/pq \\ 0 & pq/\mu \end{bmatrix},
\]

\[
B_2 = \begin{bmatrix} 0 \\ 0 \\ \mu \end{bmatrix}, \quad A_3 = \begin{bmatrix} \bar{\theta} \\ 0 \\ \mu \end{bmatrix}.
\]

Since the lower block-sub-diagonal matrix in \( \bar{P} \) is of rank 1, we could use the matrix-geometric form solution.

Let \( (Q,I) \) be the steady state of \( (Q_n, I_n) \) and its distribution is denoted as

\[
\pi_{ki} = P\{Q = k, I = i\}, \quad (k,i) \in \Omega,
\]

\[
\pi_{00} = \pi_0, \quad (\pi_{k0}, \pi_{ki}) = \pi_k, \quad k \geq 1.
\]

Due to the block tridiagonal structure of the transition probability matrix, \{\( Q_n, I_n \)\} is a quasi birth and death chain. Meanwhile, \( \bar{P} \) is a degenerate GI/M/1 type structure matrix, see Neuts [15]. It is necessary to solve for the minimal non-negative solution \( \bar{R} \), called the rate matrix, of the matrix quadratic equation.
Equilibrium Mixed Strategies in a Discrete-Time Markovian Queue

\[ \tilde{R} = \tilde{R}^2 B_1 + \tilde{R}A_1 + C_1. \]  
(3)

We obtain the explicit expression for \( \tilde{R} \) in the following theorem.

**Theorem 3.1** The matrix Equation (3) has the minimal non-negative solution

\[ \tilde{R} = \begin{bmatrix} r & \alpha \\ \bar{\alpha} & \bar{\mu} \end{bmatrix}, \]  
(4)

where \( r = \bar{\alpha} p / (\theta + \bar{\alpha} p) \) and \( \alpha = p \bar{\mu} / \bar{\mu} \).

**Proof.** Because the coefficients of Equation (3) are all upper triangular matrices, we can assume that \( \tilde{R} \) has the same structure as

\[ \tilde{R} = \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix}. \]

Substituting \( \tilde{R} \) into Equation (3), we have the following set of equations:

\[
\begin{align*}
    r_{11} &= \bar{\alpha} p r_{11} + \bar{\alpha} p, \\
    r_{12} &= (r_{11} r_{12} + r_{12} r_{22}) \bar{\mu} + r_{11} \theta \bar{\mu} + r_{12} (1 - p \bar{\mu} - \bar{\mu}) + \theta p, \\
    r_{22} &= r_{22}^2 \bar{\mu} + r_{22} (1 - p \bar{\mu} - \bar{\mu}) + p \bar{\mu}.
\end{align*}
\]
(5)

From the first equation of Equation (5), we can easily get \( r_{11}, \) denoted by \( r \). From the expression, we know \( 0 < r < 1 \). Meanwhile, the third equation of Equation (5) has two roots \( 1 \) and \( \alpha \). To obtain the minimal non-negative solution of Equation (3), take \( r_{22} = \alpha \) (<1) into the corresponding equation. Substituting \( r \) and \( \alpha \) into the second equation of Equation (5), we obtain \( r_{12} = \alpha / \bar{\mu} \).

Next, we analyze the joint probability distribution for \((Q_n, I_n)\). According to the matrix-geometric solution method (see [7]), we have

\[ \pi_k = (\pi_{k0}, \pi_{k1}) = (\pi_{10}, \pi_{11}) \tilde{R}^{k-1}, 1 \leq k \leq Q_x - 1. \]  
(6)

In addition, \((\pi_{00}, \pi_{10}, \pi_{11})\) satisfies

\[ (\pi_{00}, \pi_{10}, \pi_{11}) B[\tilde{R}] = (\pi_{00}, \pi_{10}, \pi_{11}), \]

where

\[ B[\tilde{R}] = \begin{bmatrix} A & C \\ B & RB_1 + A_1 \end{bmatrix} = \begin{bmatrix} p & \bar{\alpha} p & \theta p \\ 0 & \bar{\alpha} \bar{\mu} & \theta + \bar{\alpha} p \\ \bar{\alpha} \mu & 0 & 1 - \bar{\mu} \end{bmatrix}. \]

Taking \( \pi_{00} \) as a constant, we obtain

\[ \pi_1 = (\pi_{10}, \pi_{11}) = (r, \alpha \pi_{11}) \pi_{00}. \]

From Equation (4), we know

\[ \tilde{R}^k = \begin{bmatrix} r^k & \frac{\alpha}{\bar{\mu}} \sum_{j=0}^{k-1} r^{k-1-j} \pi_{11} \pi_{00} \\ 0 & \alpha^k \end{bmatrix}. \]
Now, substituting \( \pi_1 \) and \( \hat{\mathbf{R}} \) into Equation (6) for matrix-geometric solution and letting
\[
G = \alpha / (\bar{\mu} (r - \alpha)),
\]
we obtain
\[
\pi_k = (\pi_{k0}, \pi_{k1}) = (r^k, G(r^k - \alpha^k))\pi_{00}, \quad 1 \leq k \leq Q_v - 1.
\]
Then, we derive the stationary probabilities when \( k \) equals \( Q_v \) or \( Q_v + 1 \) from the following set of equations:
\[
\begin{align*}
\pi_{Q_v} &= \pi_{Q_v} - \pi Q_v - 1 A_2 + \pi_{Q_v + 1} B_2 = \pi_{Q_v} E_2, \\
\pi_{Q_v} &= \pi_{Q_v} + \pi_{Q_v + 1} A_3 = \pi_{Q_v + 1} E_3.
\end{align*}
\]
Therefore,
\[
\pi_{Q_v} = (\pi_{Q_v, 0}, \pi_{Q_v, 1}) = \left[ \frac{\theta + \bar{\alpha} p}{\theta + \bar{\alpha} pq} r^{Q_v}, G(r^{Q_v} - \alpha^{Q_v}) \right] \pi_{00},
\]
and
\[
\pi_{Q_v + 1} = (\pi_{Q_v + 1, 0}, \pi_{Q_v + 1, 1}) = \left[ \frac{\bar{\alpha} p q (\theta + \bar{\alpha} p)}{\theta (\theta + \bar{\alpha} pq)} r^{Q_v}, \frac{p q \left( \theta + \bar{\alpha} p \right) r^{Q_v} - \bar{G}(r^{Q_v} - \alpha^{Q_v})}{\mu} \right] \pi_{00}.
\]
The constant \( \pi_{00} \) can be determined by the normalization condition
\[
\sum_{k=0}^{Q_v} \pi_{k0} + \sum_{k=1}^{Q_v + 1} \pi_{k1} = 1.
\]

3.2. Single Vacation Policy Model

Since the analysis of the single vacation case is similar to Subsection 3.1, we only present the main results which are different from those in the multiple vacation case.

In the single vacation queue, there exists an idle state \((0, 1)\), which leads to some differences from the multiple vacation queue. The state space changes to \( \Omega = \{(k, i) \mid k \geq 0, i = 0, 1\} \) and \( \pi_0 = (\pi_{00}, \pi_{01}) \). In the transition probability matrix \( \hat{\mathbf{P}} \), now we have
\[
\mathbf{A} = \mathbf{A}_s = \begin{bmatrix} \bar{\alpha} \bar{p} & \theta \bar{p} \\ 0 & \bar{p} \end{bmatrix}, \quad \mathbf{B} = \mathbf{B}_s = \begin{bmatrix} 0 & 0 \\ \bar{p} \mu & 0 \end{bmatrix}, \quad \mathbf{C} = \mathbf{C}_s = \begin{bmatrix} \bar{\alpha} p & \theta p \\ 0 & p \end{bmatrix}.
\]
The other sub-matrixes are the same as those in multiple vacations queue.

The expressions of stationary probabilities in vacation states except \( \pi_{00} \) are the same as in multiple vacations case, i.e. the expressions of \( \pi_{k0} \) \((k \geq 1)\) in the previous section can be used for the single vacation system. The stationary probabilities for the working states are different from the multiple vacation case and are given by
\[
\begin{align*}
\pi_{01} &= \frac{\theta \bar{p}}{p} \pi_{00}, \\
\pi_{k1} &= G r^k + \left( \frac{\bar{\alpha}}{\bar{\mu} r} - G \right) \alpha^k \pi_{00}, \quad 1 \leq k \leq Q_v, \\
\pi_{Q_v + 1, 1} &= \frac{p q}{\mu} \left( \frac{\theta + \bar{\alpha} p}{\theta + \bar{\alpha} pq} + \bar{G} \right) r^{Q_v} + \left( \frac{\bar{\alpha}}{r} - \bar{G} \right) \alpha^{Q_v} \pi_{00}.
\end{align*}
\]
3.3. Equilibrium Strategies

Note that the expected benefit of an arriving customer joining a system with \( k \) customers is

\[
B = R - C(k + 1) \frac{C_i(k)}{\mu} - C \pi_{i/k}(0 | k) \frac{\theta}{\theta+\frac{\theta}{\mu}}.
\]

Also we get

\[
\pi_{i/k}(0 | k) = \frac{p_{i/k}a}{p_{i/k}a + p_{i/k}b} = \left[1 + H(k)\right]^{-1}, \quad 1 \leq k \leq Q_e - 1,
\]

(7)

\[
\pi_{i/k}(0 | Q_e) = \left[1 + \frac{a + \theta - \theta p}{\theta + \theta p} H(Q_e)\right]^{-1},
\]

(8)

\[
\pi_{i/k}(0 | Q_e + 1) = \left[1 + \frac{a + \theta - \theta p}{\theta + \theta p} H(Q_e)\right]^{-1},
\]

(9)

where

\[
H(k) = \begin{cases}
G\left(1 - \left(\frac{\alpha}{r}\right)^k\right), & k = 0, 1, 2, \ldots, \text{in multiple vacations model} \\
\theta \left(\frac{\alpha}{r}\right)^k + G\left(1 - \left(\frac{\alpha}{r}\right)^k\right), & k = 0, 1, 2, \ldots, \text{in multiple vacations model}.
\end{cases}
\]

(10)

We define the function

\[
g(k, x) = R - C(k + 1) \frac{C_i(k)}{\mu} - C \left[1 + \frac{a + \theta - \theta px}{\theta + \theta p} H(k)\right]^{-1}, \quad x \in [0, 1], \quad k = 0, 1, 2, \ldots,
\]

(11)

and use it to prove the existence of equilibrium mixed threshold strategies. We derive the corresponding thresholds.

Let

\[
g_u(k) = g(k, 1), \quad g_L(k) = g(k, 0), \quad k = 0, 1, 2, \ldots
\]

(12)

It is easy to see that in both single and multiple vacation cases \( g_u(0) > 0 \) and \( g_L(0) > 0 \). In addition, \( \lim_{k \to \infty} g_u(k) = \lim_{k \to \infty} g_L(k) = -\infty \). Hence there exists \( k_u \) such that

\[
g_u(0), g_u(1), g_u(2), \ldots, g_u(k_u) > 0 \quad \text{and} \quad g_u(k_u + 1) \leq 0.
\]

(13)

Because the function \( g(k, x) \) is increasing with respect to \( x \) for every fixed \( k \), we get the relation \( g_L(k) \leq g_u(k), \quad k = 0, 1, 2, \ldots \). In particular, \( g_L(k_u + 1) \leq 0 \) while \( g_u(0) > 0 \). Hence, there exists a \( k_L \leq k_u \) such that

\[
g_L(k_L) > 0 \quad \text{and} \quad g_L(k_L + 1), \ldots, g_L(k_u), g_L(k_u + 1) \leq 0.
\]

(14)

We can now establish the existence of the equilibrium mixed threshold strategies and give the following theorem.
Theorem 3.2 In an observable Geo/Geo/1 queue with multiple vacations or single vacation, all mixed threshold strategies \((Q_e,q_e)\) – with observed \(Q_n\), entering if \(Q_n \leq Q_e - 1\), entering with probability \(q_e\) (balking with probability \(1 - q_e\)) if \(Q_n = Q_e\), are equilibrium balking strategies and satisfy the following conditions:

If \(k_L = k_U\), there exists a unique equilibrium strategy for \(Q_e = k_L + 1, q_e = 0\);

If \(k_L < k_U\), there exists a group of equilibrium strategies for \(Q_e = k_L + 1, \ldots, k_U\), and

\[
q_e = \frac{1}{\theta} \left[ \frac{\theta + \bar{\theta} p}{H(Q_e)} \left( \frac{C}{\theta R - C(Q_e + 1)} - 1 \right) - \theta \right],
\]

(15)

where \(H(Q_e)\) is given by Equation (10).

Proof. Assuming all other customers follow the same mixed threshold strategy \((Q_e,q_e)\), we now consider a tagged customer at his arrival instant in the following cases.

(i) \(k_L = k_U\) case.

If the tagged customer finds \(k \leq Q_e - 1\) customers and decides to enter, his expected utility is equal to

\[
B = R - \frac{C(k + 1)}{\mu} - \frac{C}{\theta} \left[ 1 + H(k) \right]^{-1} = g_U(k) > 0,
\]

because of Equations (2), (7), (12) and (13). So in this case the customer prefers to enter.

If the tagged customer finds \(k = Q_e\) customers and decides to enter, his expected utility is

\[
B = R - \frac{C(Q_e + 1)}{\mu} - \frac{C}{\theta} \left[ \frac{1 + \theta + \bar{\theta} pq}{\theta + \bar{\theta} p} H(Q_e) \right]^{-1} < g_U(Q_e) \leq 0,
\]

because of Equations (2), (8), (12) and (13). Therefore in this case the customer prefers to balk.

(ii) \(k_L < k_U\) case.

Fix an \(Q_e \in \{k_L + 1, \ldots, k_U\}\) and define the corresponding \(q_e\) by Equation (15), i.e. \(q_e\) is the unique solution of \(g(Q_e, x) = 0\). The quantity \(q_e\) is a probability because \(g(Q_e, x) = 0\) is continuous with respect to \(x\) and \(g(Q_e, 0) g(Q_e, 1) = g_L(Q_e) g_U(Q_e) \leq 0\), because of Equations (13) and (14). Since the strategy \((Q_e,1)\) is equivalent to \((Q_e + 1,0)\), we will discuss \(q_e \in [0,1)\) case.

If the tagged customer finds \(k \leq Q_e - 1\) customers and decides to enter, his expected utility is

\[
B = R - \frac{C(k + 1)}{\mu} - \frac{C}{\theta} \left[ 1 + H(k) \right]^{-1} = g_U(k) > 0,
\]

because of Equations (2), (7), (12) and (13). So in this case the customer prefers to enter.
If the tagged customer finds \( k = Q_e \) customers and decides to enter, his expected utility is

\[
B = R - \frac{C(Q_e + 1)}{\mu} - \frac{C}{\theta} \left[ 1 + \frac{\theta \bar{\theta} p q_e}{\theta + \bar{\theta} p} H(Q_e) \right]^{-1} = g(Q_e) = 0,
\]

because of Equations (2), (8) and (11). So in this case the tagged customer is indifferent between entering and balking.

Thus any decision is optimal and in particular entering with probability \( q_e \) is optimal. Because if the customer chooses another \( q < q_e \), then \( B < 0 \), which could not reach an equilibrium. On the other hand, if \( q > q_e \), so \( B > 0 \) or customers prefer to enter, i.e. \( q = 1 \), which will not lead to an equilibrium either.

Figure 1. Equilibrium Indicators. Sensitivity with respect to \( R \), for \( \mu = 0.5, \theta = 0.1, p = 0.3, C = 1 \).

If the tagged customer finds \( k = Q_e + 1 \) customers and decides to enter, his expected utility is
The above argument shows that any mixed threshold strategy \((q, \bar{q})\) with \(q\) and \(\bar{q}\) satisfying the conditions of Theorem 3.2 is a best response against itself, therefore, it is an equilibrium strategy.

Because the balking probability is equal to \((p_0 + p_1)(\theta + \theta H(Q_e)) + p_0 p_1\), the social benefit per time unit when all customers follow the threshold policy given in Theorem 3.2 equals

\[
SB = \sum_{k=0}^{\infty} \left(1 - \frac{C}{\mu} \right)^{k+1} \left(1 + \frac{\theta}{\mu} + \frac{\bar{\theta}}{\mu} \frac{p_0 + p_1}{\theta + \theta H(Q_e)}\right) H(Q_e + 1),
\]

because of Equations (2), (9), (12) and (14). Therefore in this case the customer prefers to balk.

---

**Figure 2. Equilibrium Indicators. Sensitivity with respect to \(p\), for \(\mu = 0.5, \theta = 0.1, R = 12, C = 1.\)**
The above argument shows that any mixed threshold strategy \((Q_e, q_e)\) with \(Q_e\) and \(q_e\) satisfying the conditions of Theorem 3.2 is a best response against itself, therefore, it is an equilibrium strategy.

Because the balking probability is equal to \((\pi_{Q_e,0} + \pi_{Q_e,1})(1-q_e) + \pi_{Q_e+1,0} + \pi_{Q_e+1,1}\), the social benefit per time unit when all customers follow the threshold policy given in Theorem 3.2 equals

\[
SB = p[1-(\pi_{Q_e,0} + \pi_{Q_e,1})(1-q_e) - \pi_{Q_e+1,0} - \pi_{Q_e+1,1}]R - C \left[ \sum_{k=0}^{Q_e+1} k(\pi_{k0} + \pi_{k1}) \right].
\]

4. Numerical Experiments

In this section, based on the results obtained, we present a set of numerical examples to investigate the effects of system parameters and vacation policies on the equilibriums and social benefit. Each figure has four lines representing the values of the left or right equilibrium thresholds in multiple vacation queue labeled as ‘m’ and single vocation queue labeled as ‘s’. We can make the following observations.

Figure 3. Equilibrium Indicators. Sensitivity with respect to \(\theta\), for \(\mu = 0.5, p = 0.3, R = 12, C = 1\).
From the sub-figures $a_1, a_2, a_3$ in Figures 1-3, we find the thresholds increase as the parameters $R$, $p$ or $\theta$ increase. When these parameters change within certain ranges, the thresholds remain constant, showing a staircase-like increasing pattern. On the other hand, the joining probabilities fluctuate as $R$, $p$ and $\theta$ increase. Specifically, the probabilities decrease when the thresholds stay the same and go up when the thresholds increase. Especially, in a certain interval, if the threshold gets an inflection point, the probability takes an extreme value. This behavior is consistent with the reality and is easy to explain. For example, an increase in $p$ implies an increase in the arrival rate of customers. Thus, if the value of threshold is constant, the number of balking customers will increase which leads to the decrease in entrance probability. However, when the threshold goes up, this entrance probability will also jump up. Furthermore, we find that for either multiple vacation policy or single vacation policy, the entrance probability $q_e(k_v)$ is always more than $q_e(k_v+1)$. This reflects the fact that customers have a stronger inclination to join the queue when the threshold becomes larger, which belongs to the ‘Follow the Crowd’ (FTC) behavior. In addition, with a stepped increase of threshold or fluctuating rising of entrance probability, the social benefit (the benefit of both the server and customers) has a increase tendency with respect to all parameters $R$, $p$ and $\theta$. Moreover, the social benefit per unit time for the single vacation queue is obviously more than for the multiple vacation queue. Clearly, this is because we do not consider the server’s idle time cost and vacation benefit, thus the single vacation reduces the customer waiting cost, thus improving the social benefit. Otherwise, we have to re-evaluate the two vacation policies by including more cost and benefit terms in the social welfare function.

5. Conclusion

In this paper, we have studied the strategic customer behavior in the discrete-time single server queue with either multiple vacations or single vacation. Customers can decide whether to join or to balk based on the congestion and cost/benefit information. We have proved that there exist equilibrium mixed-type threshold strategies for customers which are more flexible and generalize the pure threshold strategy.

From the numerical experiments, we find that customers have a stronger inclination to enter when the threshold becomes larger, which belongs to FTC type. Moreover, it is beneficial for both the service provider and customers to increase system parameters $R$, $p$ or $\theta$ to gain more utility. Besides, the social benefit rate in a single vacation model is obviously more than that in a multiple vacation model for a simpler setting without considering the idle time cost and vacation benefit. Evaluating these systems with more complex cost and benefit structures can be a good direction of future research.

Acknowledgements

This work was supported by the National Science Foundation of China (No.11271373), China Scholarship Council and Hunan Provincial Innovation Foundation For Postgraduate (No.CX2013B042).

References


M/G/1 Multiple Vacation Model with Balking for a Class of Disciplines

Zsolt Saffer1,* and Wuyi Yue2

1Department of Networked Systems and Services, Budapest University of Technology and Economics, Budapest, Hungary
2Department of Intelligence and Informatics, Konan University, Kobe, Japan

Received November 2013, accepted May 2014

Abstract:
In this paper an M/G/1 multiple vacation model with balking is considered. In this model the probability of joining to the system upon arrival depends on the number of customers at previous (customer or vacation) start epoch. The distribution of the joining probabilities is general, i.e. it is not limited to any class of distribution s. The model enables a wide range of service disciplines including the exhaustive, the non-exhaustive, the semi-exhaustive, the gated, the G-limited and the E-limited ones. We establish stationary relationships between the number of customers in the system at different characteristic epochs. This leads to a system of linear equations for the stationary number of customers in terms of unknown probabilities. We provide the solution for these unknown probabilities on discipline specific way for all the above listed disciplines. Additionally the stability, the special case of state independent joining probabilities and the numerical solution are also discussed.

Keywords: Balking, M/G/1, queueing theory, vacation model.

1. Introduction

In a system with balking the arriving customers are allowed to decide whether they would like to join the system or not. In the later case the customer leaves the system forever. Such joining decision of an arrival customer can typically be handled by joining probability, which is alternatively also called as input probability. In such a way situations can be modeled, in which the customer is unwilling to join the system. Typical examples are e.g. customer services of banks or service providers for electricity, gas, etc.

One of the reason why customer is unwilling to join the system is that it finds too many other customers arrived and to be served before him. This can be modeled by enabling the joining probabilities to depend on the system state. Thus state dependent joining probabilities enable more realistic modeling of real-life situations. Queueing models with balking have large literature. Here we review previous works on several, more important and essentially different modeling directions.

The simplest case to incorporate balking is to have state independent joining probability. One of the early investigations of such queueing model with balking is [13], which deals with an M/1 vacation model with fix input probability. Another work on model with fix input probability is [7], in which finite buffer queue is investigated.

The situation becomes considerable more complex, when the joining probability is state dependent. A model with balking is considered in [15], in which the joining probability is state dependent and the number of customers at the system at previous (customer or vacation) epoch is taken into account. The state dependent joining probability is a more realistic modeling for real-life situation, where the customer is unwilling to join the system, because it finds too many other customers arrived and to be served before him.

* Corresponding author. E-mail: safferzs@hit.bme.hu

Authors' Biographies:

Zaiming Liu is a Professor in Central South University (CSU). He is also the dean of the School of Mathematics and Statistics at CSU. His research interests are in issues related to queueing theory and network applications, insurance risk theory, Markov process and its applications. He has published more than 60 papers in international journals such as Applied Mathematics and Computation, Nonlinear Analysis, Economic Model, etc.

Yan Ma is a Ph.D. student in the School of Mathematics and Statistics at Central South University. She is currently studying in Beedie School of Business at Simon Fraser University as a joint Ph.D. student under the support of China Scholarship Council. Her research interests include vacation queueing systems and economics of queues.

Zhe George Zhang is a Professor in Management Science at Western Washington University in the USA and a Professor in Operations Research at Simon Fraser University in Canada. He obtained his Ph.D. in Operations Research from the University of Waterloo, Canada. The focus of his current research is on theory and application of queueing and other stochastic models.