

## Optimization and sensitivity analysis of controlling arrivals in the queueing system with single working vacation

Dong-Yuh Yang<sup>a</sup>, Kuo-Hsiung Wang<sup>b,\*</sup>, Chia-Huang Wu<sup>a</sup>

<sup>a</sup> Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu 30050, Taiwan

<sup>b</sup> Department of Applied Mathematics, National Chung-Hsing University, Taichung 402, Taiwan

### ARTICLE INFO

#### Article history:

Received 2 February 2009

Received in revised form 26 November 2009

#### Keywords:

*F*-policy

Optimization

Quasi-Newton method

Sensitivity analysis

Working vacation

### ABSTRACT

This paper analyzes the *F*-policy M/M/1/*K* queueing system with working vacation and an exponential startup time. The *F*-policy deals with the issue of controlling arrivals to a queueing system, and the server requires a startup time before allowing customers to enter the system. For the queueing systems with working vacation, the server can still provide service to customers rather than completely stop the service during a vacation period. The matrix-analytic method is applied to develop the steady-state probabilities, and then obtain several system characteristics. We construct the expected cost function and formulate an optimization problem to find the minimum cost. The direct search method and Quasi-Newton method are implemented to determine the optimal system capacity *K*, the optimal threshold *F* and the optimal service rates ( $\mu_B, \mu_V$ ) at the minimum cost. A sensitivity analysis is conducted to investigate the effect of changes in the system parameters on the expected cost function. Finally, numerical examples are provided for illustration purpose.

© 2010 Elsevier B.V. All rights reserved.

### 1. Introduction

Many real-world situations involve queueing systems, in which the server may be unavailable for occasional intervals of time when the system becomes empty. The occasional intervals of time are called vacations. Queueing systems with server vacations have been investigated and applied extensively in many areas, such as computers and communication systems, manufacturing/production and inventory systems. Under various vacation policies, the optimal design and control of queues appears flexible. An excellent and comprehensive survey on this topic can be found in Doshi [1] and Takagi [2].

In this paper, we consider the *F*-policy M/M/1/*K* queueing system with single working vacation and an exponential startup time. The server takes a single vacation whenever the system becomes empty. During a vacation period, the server remains working at a different service rate rather than completely stop to provide service. Such a vacation is called a working vacation (see Servi and Finn [3]). Gupta [4] first introduced the concept of the *F*-policy which means that when the number of customers in the system reaches its capacity *K* (i.e. the system becomes full), no further arriving customers are allowed to enter the system until the queue length decreases to a certain threshold value  $F$  ( $0 \leq F \leq K - 1$ ). At that time, the server requires an exponential startup time to restart allowing customers to enter the system. This queueing system is referred to as the *F*-policy M/M/1/*K*/WV queueing system with an exponential startup time.

Past work regarding controllable queues may be divided into two parts according to whether the system is considered to control the service or the arrival. The first category of controlling the service which focuses on three different threshold policies includes the *N*-policy, introduced in [5], the *T*-policy, introduced in [6] and the *D*-policy, introduced in [7]. In

\* Corresponding author. Tel.: +886 4 22860133x509; fax: +886 4 22873028.

E-mail address: [khwang@amath.nchu.edu.tw](mailto:khwang@amath.nchu.edu.tw) (K.-H. Wang).

the second category of controlling the arrival, Gupta [4] was the first to investigate the  $F$ -policy and derived the steady-state analytic solutions. Karaesmen and Gupta [8] applied the duality relationship to obtain the stationary queue length distributions for the two queueing systems under  $N$ -policy and  $F$ -policy. Recently, Wang et al. [9] extended the  $F$ -policy  $M/M/1$  queue to the  $F$ -policy  $M/G/1$  queue by using the supplementary variable technique. Most recently, Wang et al. [10] investigated the optimal management problem of the  $F$ -policy  $G/M/1/K$  queueing systems with an exponential startup time.

Servi and Finn [3] first introduced the  $M/M/1$  queueing model with working vacation. Queueing models with working vacation have been studied by several researchers such as Kim et al. [11], Li et al. [12], Yi et al. [13], Tian et al. [14], and so on. Baba [15] examined the  $GI/M/1$  queue with multiple working vacations. Jain and Agrawal [16] investigated  $M/E_k/1$  queue with working vacation and developed the average queue length by the probability generating function. Moreover, Servi and Finn's system [3] was extended to  $M/G/1$  case in [17], where the service times during a normal busy period, the service times during working vacation, and the length of working vacation are general distributions. The finite capacity  $GI/M/1/WV$  queue with multiple working vacations was discussed in [18]. They derived the system size distributions at pre-arrival and at arbitrary epochs, the blocking probability and the mean waiting time in the system. Li and Tian [19] analyzed a discrete time  $GI/Geo/1$  queue with working vacations and presented the stochastic decomposition results for the queue length and the expected waiting time. Recently, Tian et al. [14] investigated a discrete time  $Geom/Geom/1$  queue with multiple working vacations by using the quasibirth and death process and matrix-geometric method.

To the best of our knowledge, there has been no research that explores controlling arrivals in the queueing system by using the matrix-analytic method. This motivated us to apply this approach to solve the steady-state equations of the  $F$ -policy  $M/M/1/K/WV$  queueing system with an exponential startup time. The rest of this paper is organized as follows. In the next section, some basic assumptions and the practical justification of the model are described. In Section 3, we use the matrix-analytic method to derive the closed-form expression of steady-state probability vector; and further, special cases are also discussed. Section 4 develops the various system characteristics in matrix forms. In Section 5, the expected cost function per unit time is constructed to determine the joint optimal values at the minimum cost. We employ the direct search method to find the optimal capacity,  $K^*$ , and the optimal threshold value,  $F^*$ . Subsequently, we implement the Quasi-Newton method to obtain the optimal service rates ( $\mu_B, \mu_V$ ) after  $K^*$  and  $F^*$  are determined. In Section 6, a sensitivity analysis is carried out to verify the effect of the system parameters on the cost function. Section 7 presents a numerical example to accomplish the optimum tasks. In addition, some numerical experiments are also provided to illustrate the sensitivity analysis. Finally, conclusions are given in Section 8.

## 2. Model descriptions

### 2.1. Assumptions

We investigate the  $F$ -policy  $M/M/1/K/WV$  queueing system with an exponential startup time. The basic assumptions of this controlling arrivals system is described as follows. The arrival of customer follows a Poisson process with parameter  $\lambda$ . The service times during a normal busy period are according to exponential distribution with parameter  $1/\mu_B$ . The system capability is finite, denoted by  $K$  ( $K < \infty$ ). If the system is full, it forbids any customers entering the system until the number of customer is less than or equal to a prefix threshold value  $F$ . Moreover, at this time, the server requires an exponential startup time with parameter  $\gamma$  and the customers continue enter the system. Once the system becomes empty, the server goes to a working vacation. The service times during a working vacation period and the duration of a working vacation are according to exponential distribution with parameter  $1/\mu_V$  and  $\theta$ , respectively. If the server finds no customers in the system at the end of vacation, the server waits for the next arrival, i.e., single vacation policy. Each server can serve only one customer at a time. Arrived customers form a single waiting line and follow the first-come first-served (FCFS) discipline. If all servers are busy, the arrived customers must wait in the queue until a server is available. Various stochastic processes (arrival or service or vacation) involved in this system are assumed to be independent with each other.

### 2.2. Practical justification of the model

The intention of the  $F$ -policy is to control arrival process, which focuses on reducing the number of customers in the system. In real-world applications, the model discussed in this paper is quite useful due to the consideration of arriving customers. Such a queueing model frequently occurs in the area of computer processing, transportation systems and so on. A practical problem related to a computer processing system is provided for illustration purpose. If the processor is available, indicating that it is not currently working on a task and the message is processed. Then the message is temporarily stored in a buffer to be served some time later if the processor is unavailable. When the buffer runs full at any time, newly arriving messages are blocked until the number of messages drops to a specified threshold level. Messages are immediately admitted to enter the system on condition that the system buffer reduces to the specific level. On the other hand, the processor will switch a lower processing rate to newly arriving messages whenever all messages are processed. However, the processor can switch a higher processing rate at any time. It is helpful to prevent a computer from becoming overloaded and enhance the computer performance.

### 3. Steady-state results

For the  $F$ -policy M/M/1/K/WV queueing system with an exponential startup time, we develop the steady-state probability equations based on the Markov process method.

Let us define some notations in the following:

$N(t) \equiv$  the number of customers in the system at time  $t$ ,

$Y(t) \equiv$  the state of the server at time  $t$ ,

where

$$Y(t) = \begin{cases} 0, & \text{if the arrivals are not allowed to enter the system and the server is on a working vacation period;} \\ 1, & \text{if the arrivals are not allowed to enter the system and the server is on a normal busy period;} \\ 2, & \text{if the arrivals are allowed to enter the system and the server is on a normal busy period;} \\ 3, & \text{if the arrivals are allowed to enter the system and the server is on a working vacation period.} \end{cases}$$

Then  $\{Y(t), N(t); t \geq 0\}$  is a continuous time Markov process with state space

$$S = \{(i, n) | i = 0, 1; n = 0, 1, \dots, K - 1, K\} \cup \{(i, n) | i = 2, 3; n = 0, 1, \dots, K - 2, K - 1\}.$$

The steady-state probabilities of the system are defined as follows:

$$P_i(n) = \lim_{t \rightarrow \infty} \{Y(t) = i, N(t) = n\}, \quad i = 0, 1, n = 0, 1, \dots, K - 1, K.$$

$$P_i(n) = \lim_{t \rightarrow \infty} \{Y(t) = i, N(t) = n\}, \quad i = 2, 3, n = 0, 1, \dots, K - 2, K - 1.$$

#### 3.1. Steady-state probability equations

Referring to the state-transition-rate diagram for the  $F$ -policy M/M/1/K/WV queueing system with an exponential startup time shown in Fig. 1, we have the following steady-state probability equations:

$$(\theta + \gamma)P_0(0) = \mu_V P_0(1), \tag{1}$$

$$(\mu_V + \theta + \gamma)P_0(n) = \mu_V P_0(n + 1), \quad n = 1, 2, \dots, F, \tag{2}$$

$$(\mu_V + \theta)P_0(n) = \mu_V P_0(n + 1), \quad n = F + 1, F + 2, \dots, K - 1, \tag{3}$$

$$(\mu_V + \theta)P_0(K) = \lambda P_3(K - 1), \tag{4}$$

$$\gamma P_1(0) = \mu_B P_1(1) + \theta P_0(0), \tag{5}$$

$$(\mu_B + \gamma)P_1(n) = \mu_B P_0(n + 1) + \theta P_0(n), \quad n = 1, 2, \dots, F, \tag{6}$$

$$\mu_B P_1(n) = \mu_B P_1(n + 1) + \theta P_0(n), \quad n = F + 1, F + 2, \dots, K - 1, \tag{7}$$

$$\mu_B P_1(K) = \lambda P_2(K - 1) + \theta P_0(K), \quad F \neq K - 1, \tag{8}$$

$$\lambda P_2(0) = \gamma P_1(0) + \theta P_3(0), \tag{9}$$

$$(\lambda + \mu_B)P_2(n) = \gamma P_1(n) + \lambda P_2(n - 1) + \mu_B P_2(n + 1) + \theta P_3(n), \quad n = 1, 2, \dots, F - 1, F, \tag{10}$$

$$(\lambda + \mu_B)P_2(n) = \lambda P_2(n - 1) + \mu_B P_2(n + 1) + \theta P_3(n), \quad n = F + 1, F + 2, \dots, K - 2, \tag{11}$$

$$(\lambda + \mu_B)P_2(K - 1) = \lambda P_2(K - 2) + \theta P_3(K - 1), \tag{12}$$

$$(\lambda + \theta)P_3(0) = \gamma P_0(0) + \mu_B P_2(1) + \mu_V P_3(1), \tag{13}$$

$$(\lambda + \mu_V + \theta)P_3(n) = \gamma P_0(n) + \lambda P_3(n - 1) + \mu_V P_3(n + 1), \quad n = 1, 2, \dots, F, \tag{14}$$

$$(\lambda + \mu_V + \theta)P_3(n) = \lambda P_3(n - 1) + \mu_V P_3(n + 1), \quad n = F + 1, F + 2, \dots, K - 2, \tag{15}$$

$$(\lambda + \mu_V + \theta)P_3(K - 1) = \lambda P_3(K - 2). \tag{16}$$

#### 3.2. Matrix form equations

Using the matrix-analytic method, we may develop the steady-state probabilities for the  $F$ -policy M/M/1/WV queueing system with an exponential startup time. The corresponding transition rate matrix  $Q$  of this Markov chain has the block-

Server State

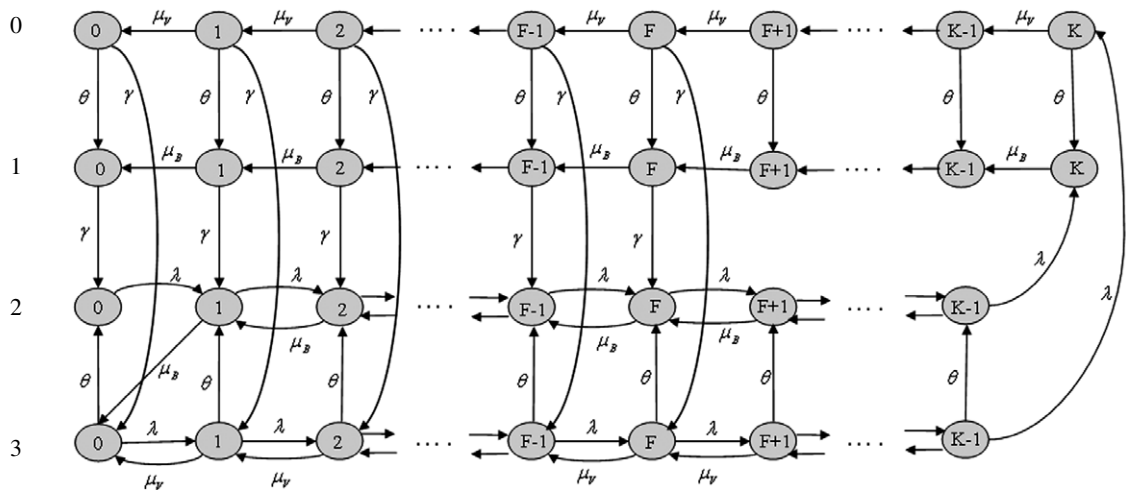


Fig. 1. The state-transition-rate diagram for an F-policy M/M/1/WV queueing system with an exponential startup time.

tridiagonal structure:

$$\mathbf{Q} = \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ F \\ F+1 \\ F+2 \\ \vdots \\ K-2 \\ K-1 \\ K \end{matrix} \begin{bmatrix} A_0 & C_0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ B_0 & A_1 & C_0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & B_1 & A_1 & C_0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & B_1 & A_1 & C_0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & B_1 & A_2 & C_0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & B_1 & A_2 & C_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & B_1 & A_2 & C_0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & B_1 & A_2 & C_1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & B_2 & A_3 \end{bmatrix},$$

where the lower boundary block entries  $A_3, B_2$  and  $C_1$  are matrices with dimensions  $(2 \times 2)$ ,  $(2 \times 4)$  and  $(4 \times 2)$ , respectively. Other matrices are square matrices with dimension  $(4 \times 4)$ . Each entry of the matrix  $\mathbf{Q}$  is listed in the following:

$$A_j = \begin{bmatrix} -(\Omega\{j \neq 0\}\mu_V + \theta + \Omega\{j \neq 2\}\gamma) & \theta & 0 & \Omega\{j \neq 2\}\gamma \\ 0 & -(\Omega\{j \neq 0\}\mu_B + \Omega\{j \neq 2\}\gamma) & \Omega\{j \neq 2\}\gamma & 0 \\ 0 & 0 & -(\lambda + \Omega\{j \neq 0\}\mu_B) & 0 \\ 0 & 0 & \theta & -(\lambda + \Omega\{j \neq 0\}\mu_V + \theta) \end{bmatrix},$$

where  $j = 0, 1, 2$  and  $\Omega\{\alpha\} = \begin{cases} 1, & \text{if } \alpha \text{ is true,} \\ 0, & \text{if } \alpha \text{ is false.} \end{cases}$

$$A_3 = \begin{bmatrix} -(\mu_V + \theta) & \theta \\ 0 & -\mu_B \end{bmatrix}, \quad B_j = \begin{bmatrix} \mu_V & 0 & 0 & 0 \\ 0 & \mu_B & 0 & 0 \\ 0 & 0 & \Omega\{j = 1\}\mu_B & \Omega\{j = 0\}\mu_B \\ 0 & 0 & 0 & \mu_V \end{bmatrix}, \quad \text{where } j = 0, 1.$$

$$B_2 = \begin{bmatrix} \mu_V & 0 & 0 & 0 \\ 0 & \mu_B & 0 & 0 \end{bmatrix}, \quad C_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{bmatrix}, \quad C_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \lambda \\ \lambda & 0 \end{bmatrix}.$$

The steady-state probability vector  $\mathbf{P}$  for  $\mathbf{Q}$  is partitioned as  $(\mathbf{P}_0, \mathbf{P}_1, \dots, \mathbf{P}_{K-1}, \mathbf{P}_K)$ , where the sub-vectors  $\mathbf{P}_n = \{P_0(n), P_1(n), P_2(n), P_3(n)\} (0 \leq n \leq K - 1)$  and  $\mathbf{P}_K = \{P_0(K), P_1(K)\}$  are of dimensions four and two, respectively. Solving the steady-state probability equations is equivalent to solve  $\mathbf{PQ} = \mathbf{0}$  along with the boundary condition  $\mathbf{P}e = 1$ , where  $e$  is a

column vector of ones. Thus, the steady-state probability equations in matrix form are expressed as follows:

$$\mathbf{P}_0 A_0 + \mathbf{P}_1 B_0 = \mathbf{0}, \tag{17}$$

$$\mathbf{P}_{n-1} C_0 + \mathbf{P}_n A_1 + \mathbf{P}_{n+1} B_1 = \mathbf{0}, \quad 1 \leq n \leq F, \tag{18}$$

$$\mathbf{P}_{n-1} C_0 + \mathbf{P}_n A_2 + \mathbf{P}_{n+1} B_1 = \mathbf{0}, \quad F + 1 \leq n \leq K - 2, \tag{19}$$

$$\mathbf{P}_{K-2} C_0 + \mathbf{P}_{K-1} A_2 + \mathbf{P}_K B_2 = \mathbf{0}, \tag{20}$$

$$\mathbf{P}_{K-1} C_1 + \mathbf{P}_K A_3 = \mathbf{0}. \tag{21}$$

### 3.3. Computation of the steady-state solution

In the following, we derive the steady-state probabilities in the matrix form by simple algebraic manipulation. Since  $A_0$  is nonsingular, we have from (17) that

$$\mathbf{P}_0 = \mathbf{P}_1 X_0, \quad \text{where } X_0 = -B_0 A_0^{-1}. \tag{22}$$

From (18) and (22), it implies that

$$\mathbf{P}_n = \mathbf{P}_{n+1} X_n, \quad \text{where } X_n = -B_1 (X_{n-1} C_0 + A_1)^{-1}, \quad 1 \leq n \leq F. \tag{23}$$

Using (19) and (23), it leads the following result,

$$\mathbf{P}_n = \mathbf{P}_{n+1} X_n, \quad \text{where } X_n = -B_1 (X_{n-1} C_0 + A_2)^{-1}, \quad F + 1 \leq n \leq K - 2. \tag{24}$$

Finally, from (20) and (24), we get

$$\mathbf{P}_{K-1} = \mathbf{P}_K X_{K-1}, \quad \text{where } X_{K-1} = -B_2 (X_{K-2} C_0 + A_2)^{-1}. \tag{25}$$

Solving (22)–(25) recursively, the solution  $\mathbf{P}_n$  ( $0 \leq n \leq K - 1$ ) can be represented in terms of  $\mathbf{P}_K$ .

$$\mathbf{P}_n = \mathbf{P}_{n+1} X_n = \cdots = \mathbf{P}_K \prod_{\xi=1}^{K-n} X_{K-\xi} = \mathbf{P}_K \Psi_n^*, \tag{26}$$

where  $\Psi_n^* = \prod_{\xi=1}^{K-n} X_{K-\xi}$  and  $X_n$  ( $0 \leq n \leq K - 1$ ) are given in (22)–(25).

From the normalizing condition and (26), we have

$$\begin{aligned} \sum_{n=0}^{K-1} \mathbf{P}_n e_1 + \mathbf{P}_K e_2 &= [\mathbf{P}_0 + \mathbf{P}_1 + \cdots + \mathbf{P}_{K-1}] e_1 + \mathbf{P}_K e_2 \\ &= [\mathbf{P}_K \Psi_0^* + \mathbf{P}_K \Psi_1^* + \cdots + \mathbf{P}_K \Psi_{K-1}^*] e_1 + \mathbf{P}_K e_2 \\ &= \mathbf{P}_K \left[ \sum_{n=0}^{K-1} \Psi_n^* e_1 + e_2 \right] = 1, \end{aligned} \tag{27}$$

where  $e_1$  and  $e_2$  are column vectors of orders four and two, respectively. All their elements are equal to one.

Note that (21) can be written as

$$\mathbf{P}_K [X_{K-1} C_1 + A_3] = \mathbf{0}. \tag{28}$$

Therefore,  $\mathbf{P}_K$  can be obtained by solving (27) and (28). Once  $\mathbf{P}_K$  have been determined, it is possible to obtain the steady-state solutions for  $\mathbf{P}_n$  ( $0 \leq n \leq K - 1$ ) from (26). To this end, a computer program (MAPLE) is developed to compute  $\mathbf{P}_n$  ( $0 \leq n \leq K$ ).

### 3.4. Special cases

We present three special cases of our model in the following.

Case 1: As  $\mu_V = 0$ , our model can be reduced to the  $F$ -policy M/M/1/K queueing system with single vacation.

Case 2: The corresponding result for the  $F$ -policy M/M/1/K queueing system by setting  $\theta$  to approach  $\infty$  and  $\mu_V = \mu_B$ , which is in accordance with the result of Gupta [4].

Case 3: The server is always on working vacation under  $F$ -policy if we set  $\theta = 0$ .

#### 4. System characteristics

Our analysis is based on the following system characteristics of the  $F$ -policy M/M/1/K/WV queueing system with an exponential startup time. Let us define

$L_S \equiv$  the expected number of customers in the system;

$P_B \equiv$  the probability that the server is busy;

$P_S \equiv$  the probability that the server starts to allow customers entering the system;

$P_L \equiv$  the probability that the system is blocked;

$W_S \equiv$  the expected waiting time in the system;

$\lambda_{eff} \equiv$  the effective arrival rate.

The expressions for  $L_S$ ,  $P_B$ ,  $P_S$ ,  $P_L$ ,  $W_S$  and  $\lambda_{eff}$  are given by:

$$L_S = \mathbf{P}\boldsymbol{\pi} = \mathbf{P}_K \left( \sum_{n=1}^{K-1} n\Psi_n^* e_1 + Ke_2 \right), \quad (29)$$

$$P_B = \sum_{n=1}^{K-1} \mathbf{P}_n e_1 + \mathbf{P}_K e_2 = 1 - P_K \Psi_0^* e_1, \quad (30)$$

$$P_S = \sum_{n=0}^F \mathbf{P}_n v_1 = \mathbf{P}_K \sum_{n=0}^F \Psi_n^* v_1, \quad (31)$$

$$P_L = \sum_{n=0}^{K-1} \mathbf{P}_n v_1 + \mathbf{P}_K e_2 = \mathbf{P}_K \left( \sum_{n=0}^{K-1} \Psi_n^* v_1 + e_2 \right), \quad (32)$$

$$W_S = \frac{L_S}{\lambda_{eff}} = \frac{\mathbf{P}_K \left( \sum_{n=1}^{K-1} n\Psi_n^* e_1 + Ke_2 \right)}{\lambda \mathbf{P}_K \sum_{n=0}^{K-1} \Psi_n^* v_2}, \quad (33)$$

$$\lambda_{eff} = \lambda \sum_{n=0}^{K-1} \mathbf{P}_n v_2 = \lambda \mathbf{P}_K \sum_{n=0}^{K-1} \Psi_n^* v_2, \quad (34)$$

where  $\mathbf{P} = (\mathbf{P}_0, \mathbf{P}_1, \dots, \mathbf{P}_{K-1}, \mathbf{P}_K)$ ,  $\boldsymbol{\pi} = (\pi_0, \pi_1, \dots, \pi_{K-1}, \pi_K)^T$ ,  $\pi_n = ne_1$  for  $0 \leq n \leq K-1$ ,  $\pi_K = Ke_2$ ,  $v_1 = [1 \ 1 \ 0 \ 0]^T$ ,  $v_2 = [0 \ 0 \ 1 \ 1]^T$ .

#### 5. Cost optimization analysis

We develop the expected cost function per unit time for the  $F$ -policy M/M/1/K/WV queue with an exponential startup time. In our cost function, four decision variables  $K$ ,  $F$ ,  $\mu_B$  and  $\mu_V$  are considered. Our objective is to determine the optimal threshold value  $F$ , say  $F^*$ , the optimal capacity  $K$ , say  $K^*$ , the optimal service rate during a normal busy period  $\mu_B$ , say  $\mu_B^*$ , and the optimal service rate during a working vacation period  $\mu_V$ , say  $\mu_V^*$ . The decision maker would implement those optimal values to minimize the expected cost per unit time. Let us define the following cost elements:

$C_h \equiv$  holding cost per unit time for each customer present in the system;

$C_b \equiv$  cost per unit time when the server is busy;

$C_l \equiv$  fixed cost for every lost customer when the system is blocked;

$C_s \equiv$  startup cost per unit time for allowing customer to enter the system;

$C_w \equiv$  waiting cost per unit time when one customer is waiting for service;

$C_k \equiv$  the fixed cost for the system capacity;

$C_1 \equiv$  cost per unit time for service during a normal busy period;

$C_2 \equiv$  cost per unit time for service during a working vacation period.

Utilizing the definitions of each cost element listed above and its corresponding system characteristics, the expected cost function per unit time is given by

$$TC(F, K, \mu_B, \mu_V) = C_h L_S + C_b P_B + C_l \lambda P_L + C_s P_S + C_w W_S + C_k K + C_1 \mu_B + C_2 \mu_V. \quad (35)$$

The cost elements in (35) are assumed to be linear in the expected number of the indicated quantity. Substitution of (29)–(34) into (35), the cost function  $TC(F, K, \mu_B, \mu_V)$  is too detailed to be shown here. As a result it would have been an arduous task to develop optimal solution  $(F^*, K^*, \mu_B^*, \mu_V^*)$  analytically, due to the highly nonlinear and complex nature of the optimization problem. We will present the numerical experiments to show that the cost function is indeed convex and that the solution gives a minimum. First, we use direct search method to find the joint optimal values  $(F^*, K^*)$  when  $\mu_B$  and

$\mu_V$  are fixed. Subsequently, we fix  $(F^*, K^*)$  and apply Quasi-Newton method to adjust the service rates  $(\mu_B, \mu_V)$  until the minimum  $TC(F^*, K^*, \mu_B, \mu_V)$  is achieved, say  $TC(F^*, K^*, \mu_B^*, \mu_V^*)$ .

5.1. Direct search method

As mentioned above, the formula (35) can be found to be too long and complicated. It is rather difficult to develop analytic results for the optimal value. Additionally, it is not easy to show convexity of this function. In order to guarantee that the optimum is obtained in the desired region, we use the direct search method to find joint optimal values,  $(F^*, K^*)$ . An efficient and direct procedure in the following is proposed for achieving  $(F^*, K^*)$ :

- Step 1: Find the optimal system capacity  $K^*$ , for the threshold value  $F$ , i.e.,  $Minimize_K TC(F, K) = TC(F, K^*)$ .
- Step 2: Compute the set of all minimum cost solutions for  $F = 0, 1, \dots, K - 1$ , i.e.,  $\Theta = \{TC(F, K^*) | F = 0, 1, \dots, K - 2, K - 1\}$ .
- Step 3: Determine the optimal operating  $F$ -policy,  $F^*$ , i.e.,  $Minimize_F \Theta = TC(F^*, K^*)$ .

5.2. The Quasi-Newton method

The solution corresponding the minimum expected cost  $TC(F^*, K^*, \mu_B, \mu_V)$ , is denoted by  $(F^*, K^*, \mu_B^*, \mu_V^*)$ . We can further decrease the minimum expected cost  $TC(F^*, K^*, \mu_B, \mu_V)$  easily by adjusting  $\mu_B$  and  $\mu_V$ . After the determination of  $(F^*, K^*)$ , we will use Quasi-Newton method to globally search  $(\mu_B, \mu_V)$  until the minimum value of  $TC(F^*, K^*, \mu_B, \mu_V)$  is achieved. The optimum problem can be illustrated mathematically as follows:

$$TC(F^*, K^*, \mu_B^*, \mu_V^*) = Minimize_{\mu_B, \mu_V} TC(F^*, K^*, \mu_B, \mu_V). \tag{36}$$

Quasi-Newton method is reliable and efficient for finding a minimizer of a nonlinear function. For subsequent iteration, this method is used to decide a search direction. Then trying different step length along this direction for a better solution until the tolerance is acceptable. We designate the vector  $\vec{\Omega}$  consisting of  $\mu_B$  and  $\mu_V$ , and construct the respective gradient  $\vec{\nabla} TC(\vec{\Omega}_0)$  which consists of  $\partial TC / \partial \mu_B$  and  $\partial TC / \partial \mu_V$ . Let the corresponding solution be denoted by  $(\mu_B^*, \mu_V^*)$ . In order to use the Quasi-Newton method conveniently, a step-by-step procedure is provided as below.

- Step 1: Let  $\vec{\Omega}_0 = [\mu_B, \mu_V]^T, i = 0$  and the tolerance  $\varepsilon = 10^{-7}$ .
- Step 2: Set the initial trial solution for  $\vec{\Omega}_0$ , and compute  $TC(\vec{\Omega}_0)$ .
- Step 3: Compute the cost gradient  $\vec{\nabla} TC(\vec{\Omega}_i) = [\partial TC / \partial \mu_B, \partial TC / \partial \mu_V]^T |_{\vec{\Omega}_i}$  and the cost Hessian matrix

$$H(\vec{\Omega}_i) = \begin{bmatrix} \partial^2 TC / \partial \mu_B^2 & \partial^2 TC / \partial \mu_B \partial \mu_V \\ \partial^2 TC / \partial \mu_V \partial \mu_B & \partial^2 TC / \partial \mu_V^2 \end{bmatrix} \text{ at the point } \vec{\Omega}_i.$$

- Step 4. Find the new trial solution  $\vec{\Omega}_{i+1} = \vec{\Omega}_i - [H(\vec{\Omega}_i)]^{-1} \vec{\nabla} TC(\vec{\Omega}_i)$ .
- Step 5. Set  $i = i + 1$  and repeat steps 3–4 until  $\text{Max}(|\partial TC / \partial \mu_B|, |\partial TC / \partial \mu_V|) < \varepsilon$ .
- Step 6. Find the global minimum value  $TC(\mu_B^*, \mu_V^*) = TC(\vec{\Omega}_i^*)$ .

6. Sensitivity analysis for the expected cost function

With the developed cost function, we perform a sensitivity analysis for the expected cost function with respect to changes in specific values of the system parameters. Differentiating  $\mathbf{PQ} = \mathbf{0}$  with respect to  $\lambda$ , we obtain

$$\frac{\partial \mathbf{P}}{\partial \lambda} \mathbf{Q} + \mathbf{P} \frac{\partial \mathbf{Q}}{\partial \lambda} = \mathbf{0}, \tag{37}$$

or equivalently

$$\frac{\partial \mathbf{P}}{\partial \lambda} = -\mathbf{P} \frac{\partial \mathbf{Q}}{\partial \lambda} \mathbf{Q}^{-1}. \tag{38}$$

We have the solutions  $\partial P_i(n) / \partial \lambda$  and  $\partial P_j(K) / \partial \lambda$  from (38) for  $0 \leq i \leq 3, 0 \leq n \leq K - 1$  and  $j = 0, 1$ . Using the same procedure,  $\partial \mathbf{P} / \partial \mu_B, \partial \mathbf{P} / \partial \mu_V, \partial \mathbf{P} / \partial \theta$  and  $\partial \mathbf{P} / \partial \gamma$  can be obtained.

Next, we differentiate the expected cost function in (35) with respect to  $\lambda$ . The sensitivity of the expected cost function is calculated as follows:

$$\frac{\partial TC}{\partial \lambda} = C_h \times \frac{\partial L_S}{\partial \lambda} + C_b \times \frac{\partial P_B}{\partial \lambda} + C_l \times \left( P_L + \lambda \frac{\partial P_L}{\partial \lambda} \right) + C_s \times \frac{\partial P_S}{\partial \lambda} + C_w \times \frac{\partial W_S}{\partial \lambda}. \tag{39}$$

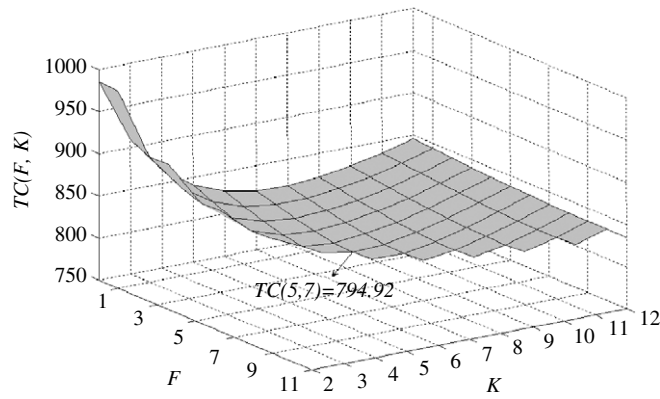


Fig. 2. The expected cost  $TC(F, K)$  for different values of  $F$  and  $K$ .

Table 1

The expected cost  $TC(F, K)$  for  $\lambda = 4.0, \mu_B = 5.0, \mu_V = 3.0, \gamma = 2.0$  and  $\theta = 3.0$ .

| K  | F      |        |        |        |        |        |        |        |        |        |        |        |   |
|----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---|
|    | 0      | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      | 10     | 11     |   |
| 2  | 985.60 | 983.54 | -      | -      | -      | -      | -      | -      | -      | -      | -      | -      | - |
| 3  | 910.04 | 898.72 | 899.38 | -      | -      | -      | -      | -      | -      | -      | -      | -      | - |
| 4  | 867.11 | 855.08 | 846.06 | 847.06 | -      | -      | -      | -      | -      | -      | -      | -      | - |
| 5  | 841.32 | 831.00 | 821.54 | 815.53 | 817.00 | -      | -      | -      | -      | -      | -      | -      | - |
| 6  | 826.69 | 818.38 | 810.32 | 803.81 | 800.15 | 801.89 | -      | -      | -      | -      | -      | -      | - |
| 7  | 819.88 | 813.35 | 806.87 | 801.23 | 797.01 | 794.92 | 796.73 | -      | -      | -      | -      | -      | - |
| 8  | 818.70 | 813.62 | 808.53 | 803.94 | 800.18 | 797.53 | 796.45 | 798.19 | -      | -      | -      | -      | - |
| 9  | 821.65 | 817.70 | 813.73 | 810.09 | 806.98 | 804.54 | 802.93 | 802.48 | 804.08 | -      | -      | -      | - |
| 10 | 827.64 | 824.58 | 821.49 | 818.63 | 816.13 | 814.07 | 812.53 | 811.61 | 811.53 | 812.97 | -      | -      | - |
| 11 | 835.91 | 833.53 | 831.13 | 828.88 | 826.90 | 825.22 | 823.90 | 822.96 | 822.48 | 822.62 | 823.89 | -      | - |
| 12 | 845.88 | 844.03 | 842.16 | 840.40 | 838.83 | 837.49 | 836.40 | 835.57 | 835.04 | 834.84 | 835.10 | 836.21 | - |

Differentiating the expected cost function in (35) with respect to  $\mu_B, \mu_V, \gamma$  and  $\theta$ , respectively, it follows that

$$\frac{\partial TC}{\partial \phi} = C_h \times \frac{\partial L_S}{\partial \phi} + C_b \times \frac{\partial P_B}{\partial \phi} + \lambda C_l \times \frac{\partial P_L}{\partial \phi} + C_s \times \frac{\partial P_S}{\partial \phi} + C_w \times \frac{\partial W_S}{\partial \phi} + \delta, \tag{40}$$

where  $\phi = \mu_B, \mu_V, \gamma, \theta$  and  $\delta = \begin{cases} C_1, & \text{if } \phi = \mu_B, \\ C_2, & \text{if } \phi = \mu_V, \\ 0, & \text{if } \phi = \gamma, \theta. \end{cases}$

### 7. Numerical examples

First, we present a numerical example for finding the joint optimal values  $(F^*, K^*)$  by using the direct search method. The following cost elements are considered:  $C_h = \$5/\text{unit}, C_b = \$300/\text{day}, C_l = \$200/\text{day}, C_s = \$400/\text{day}, C_w = \$60/\text{day}, C_k = \$15/\text{unit}, C_1 = 50/\text{unit}$  and  $C_2 = 20/\text{unit}$ . We fix  $\lambda = 4.0, \mu_B = 5.0, \mu_V = 3.0, \gamma = 2.0, \theta = 3.0$ , vary the threshold value  $F$  from 0 to  $K - 1$ , and  $K$  ranges from 2 to 12.

If the function  $TC(F, K)$  is unimodal, a single relative minimum exists. In order to find  $(F^*, K^*)$ , we should show the existence of convexity or unimodality of  $TC(F, K)$ . The curve representing the expected cost function is shown in Fig. 2. As can be seen Fig. 2, it convinces us that the expected cost function is convex. The expected cost  $TC(F, K)$  is summarized in Table 1 for various values of  $F$  and  $K$ . Using the direct search procedure, we can find that the minimum expected cost per day of \$794.920 is obtained with  $F^* = 5$  and  $K^* = 7$ .

#### 7.1. Sensitivity analysis for $F^*$ and $K^*$

Next, we perform a sensitivity analysis on the optimal threshold value  $(F^*)$  and optimal capacity  $(K^*)$  based on changes in the specific values of the system parameter. The optimal values,  $F^*, K^*$  and the corresponding minimum expected cost  $TC(F^*, K^*)$  are shown in Tables 2–4 for the following three cases, respectively.

Case 1:  $\mu_B = 5.0, \mu_V = 3.0, \gamma = 3.0, \theta = 2.0$  for different values of  $\lambda$ .

Case 2:  $\lambda = 4.0, \gamma = 3.0, \theta = 2.0$  for different values of  $(\mu_B, \mu_V)$ .

Case 3:  $\lambda = 4.0, \mu_B = 5.0, \mu_V = 3.0$  for different values of  $(\theta, \gamma)$ .



**Table 2**  
Optimal values of  $F^*$ ,  $K^*$  and the minimum expected cost for different values of  $\lambda$  ( $\mu_B = 5.0$ ,  $\mu_V = 3.0$ ,  $\gamma = 3.0$ ,  $\theta = 2.0$ ).

| $\lambda$      | 1.0     | 2.0     | 3.0     | 4.0     | 5.0     |
|----------------|---------|---------|---------|---------|---------|
| $F^*$          | 2       | 3       | 4       | 5       | 5       |
| $K^*$          | 3       | 4       | 6       | 7       | 8       |
| $TC(F^*, K^*)$ | 446.861 | 553.542 | 667.539 | 794.920 | 950.423 |

**Table 3**  
Optimal values of  $F^*$ ,  $K^*$  and the minimum expected cost for different values of  $(\mu_B, \mu_V)$  ( $\lambda = 4.0$ ,  $\gamma = 3.0$ ,  $\theta = 2.0$ ).

| $(\mu_B, \mu_V)$ | (2.0, 3.0) | (4.0, 3.0) | (6.0, 3.0) | (5.0, 1.0) | (5.0, 3.0) | (5.0, 6.0) |
|------------------|------------|------------|------------|------------|------------|------------|
| $F^*$            | 0          | 3          | 6          | 6          | 5          | 4          |
| $K^*$            | 3          | 6          | 7          | 8          | 7          | 6          |
| $TC(F^*, K^*)$   | 1026.323   | 847.979    | 784.025    | 800.806    | 794.920    | 801.146    |

**Table 4**  
Optimal values of  $F^*$ ,  $K^*$  and the minimum expected cost for different values of  $(\gamma, \theta)$  ( $\lambda = 4.0$ ,  $\mu_B = 5.0$ ,  $\mu_V = 3.0$ ).

| $(\gamma, \theta)$ | (1.0, 2.0) | (3.0, 2.0) | (6.0, 2.0) | (3.0, 0.5) | (3.0, 1.0) | (3.0, 3.0) |
|--------------------|------------|------------|------------|------------|------------|------------|
| $F^*$              | 7          | 5          | 4          | 6          | 6          | 5          |
| $K^*$              | 9          | 7          | 6          | 8          | 8          | 7          |
| $TC(F^*, K^*)$     | 830.281    | 794.920    | 781.249    | 875.276    | 828.551    | 780.993    |

**Table 5**  
Quasi-Newton method in searching the optimal solution  $\mu_B^*$  and  $\mu_V^*$  with  $\lambda = 4.0$ ,  $\gamma = 3.0$  and  $\theta = 2.0$ .

| No. of iterations        | 0       | 1       | 2       | 3       | 4       |
|--------------------------|---------|---------|---------|---------|---------|
| $TC(F, K, \mu_B, \mu_V)$ | 794.920 | 781.993 | 781.524 | 781.523 | 781.523 |
| $(F^*, K^*)$             | (5, 7)  | (5, 7)  | (5, 7)  | (5, 7)  | (5, 7)  |
| $\mu_B$                  | 5.0000  | 5.7581  | 5.8826  | 5.8891  | 5.8891* |
| $\mu_V$                  | 3.0000  | 4.4526  | 4.1218  | 4.1318  | 4.1318* |

**Table 6**  
Quasi-Newton method in searching the optimal solution  $\mu_B^*$  and  $\mu_V^*$  with  $\mu_B = 5.0$ ,  $\mu_V = 3.0$ ,  $\gamma = 3.0$ ,  $\theta = 2.0$  and for various values of  $\lambda$ .

| $\lambda$                        | 1.0        | 2.0        | 3.0        | 4.0        | 5.0        |
|----------------------------------|------------|------------|------------|------------|------------|
| $(F^*, K^*)$                     | (2, 3)     | (3, 4)     | (4, 6)     | (5, 7)     | (5, 8)     |
| $(\mu_B, \mu_V)$                 | (5.0, 3.0) | (5.0, 3.0) | (5.0, 3.0) | (5.0, 3.0) | (5.0, 3.0) |
| $TC(F^*, K^*, \mu_B, \mu_V)$     | 446.861    | 553.542    | 667.539    | 794.920    | 950.423    |
| $\mu_B^*$                        | 2.831      | 3.902      | 4.872      | 5.889      | 6.914      |
| $\mu_V^*$                        | 0.769      | 2.456      | 3.257      | 4.132      | 4.992      |
| $TC(F^*, K^*, \mu_B^*, \mu_V^*)$ | 379.884    | 538.945    | 667.125    | 781.523    | 890.685    |

One can easily see from Table 2 that (i)  $F^*$  and  $K^*$  increase as  $\lambda$  increases; and (ii) the minimum expected cost  $TC(F^*, K^*)$  increases as  $\lambda$  increases. From Table 3, it appears that (i)  $F^*$  and  $K^*$  increase as  $\mu_B$  increases or  $\mu_V$  decreases; (ii)  $TC(F^*, K^*)$  decreases as  $\mu_B$  increases; and (iii)  $TC(F^*, K^*)$  decreases or increases as  $\mu_V$  changes. From Table 4, we observe that (i)  $F^*$  and  $K^*$  decrease as  $\gamma$  or  $\theta$  increases; and (ii)  $TC(F^*, K^*)$  decreases as  $\gamma$  or  $\theta$  increases.

7.2. Searching the optimal values of  $(\mu_B, \mu_V)$

After the determination of  $F^*$  and  $K^*$ , we then apply the procedure of the Quasi-Newton method to search the optimal service rates  $(\mu_B, \mu_V)$ . Let us take the same example for illustration as above, the initial trial solution is  $(F^*, K^*, \mu_B, \mu_V) = (5, 7, 5.0, 3.0)$  with the initial value \$794.920. It can be seen from Table 5 that after only four iterations, the minimum expected cost per day of \$781.523 is achieved at  $(F^*, K^*, \mu_B^*, \mu_V^*) = (5, 7, 5.8891, 4.1318)$ . The minimum expected cost is approximately 1.8% lower from the initial value. Thus, it leads to the conclusion that the Quasi-Newton method is working well and converges very fast. In addition, we also provide other numerical results by using the Quasi-Newton method from Tables 2–4 and summarized in Tables 6–8. From Tables 6–8, it is obvious that the expected cost can be reduced essentially.

7.3. Sensitivity analysis for the expected cost function

Finally, we will use a graphical analysis to study the effects of various parameters on the expected cost function. We fix  $K = 12$ , choose  $F = 3, 6, 9$  and consider the following five cases.

- Case 4:  $\mu_B = 5.0$ ,  $\mu_V = 3.0$ ,  $\gamma = 3.0$ ,  $\theta = 2.0$ ,  $\lambda$  varies from 2.0 to 4.0.
- Case 5:  $\lambda = 4.0$ ,  $\mu_V = 3.0$ ,  $\gamma = 3.0$ ,  $\theta = 2.0$ ,  $\mu_B$  varies from 4.0 to 6.0.
- Case 6:  $\lambda = 4.0$ ,  $\mu_B = 5.0$ ,  $\gamma = 3.0$ ,  $\theta = 2.0$ ,  $\mu_V$  varies from 2.0 to 4.0.

**Table 7**

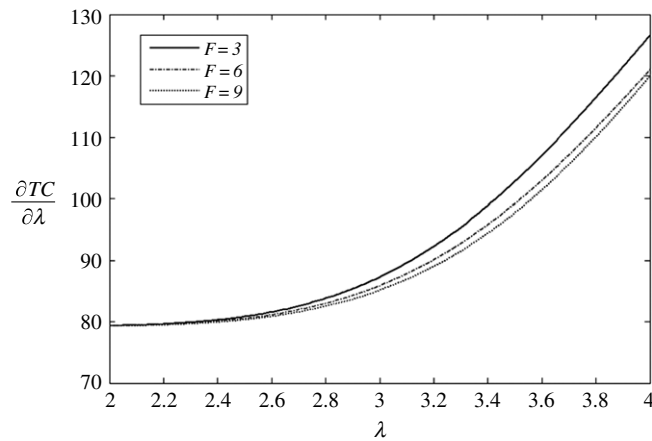
Quasi-Newton method in searching the optimal solution  $\mu_B^*$  and  $\mu_V^*$  with  $\lambda = 4.0, \gamma = 3.0, \theta = 2.0$  and for various values of  $(\mu_B, \mu_V)$ .

| $(\mu_B, \mu_V)$                 | (2.0, 3.0) | (4.0, 3.0) | (6.0, 3.0) | (5.0, 1.0) | (5.0, 3.0) | (5.0, 6.0) |
|----------------------------------|------------|------------|------------|------------|------------|------------|
| $(F^*, K^*)$                     | (0, 3)     | (3, 6)     | (6, 7)     | (6, 8)     | (5, 7)     | (4, 6)     |
| $TC(F^*, K^*, \mu_B, \mu_V)$     | 1026.323   | 847.979    | 784.025    | 800.806    | 794.920    | 801.146    |
| $\mu_B^*$                        | 5.386      | 5.928      | 5.928      | 5.837      | 5.889      | 5.927      |
| $\mu_V^*$                        | 5.523      | 4.892      | 4.122      | 3.512      | 4.132      | 4.772      |
| $TC(F^*, K^*, \mu_B^*, \mu_V^*)$ | 900.282    | 787.466    | 781.511    | 784.321    | 781.523    | 784.148    |

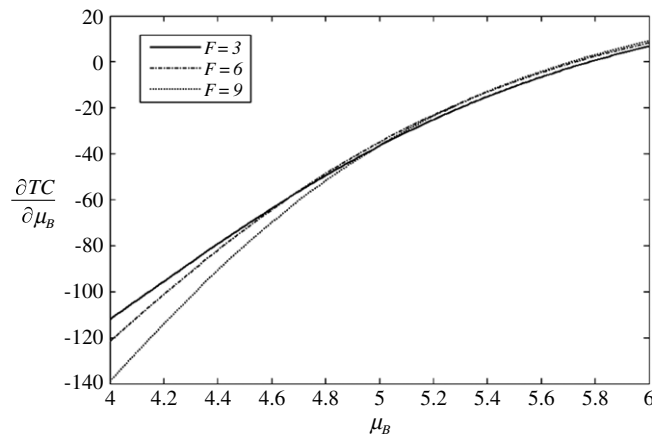
**Table 8**

Quasi-Newton method in searching the optimal solution  $\mu_B^*$  and  $\mu_V^*$  with  $\lambda = 4.0, \mu_B = 5.0, \mu_V = 3.0$  and for various values of  $(\gamma, \theta)$ .

| $(\gamma, \theta)$               | (1.0, 2.0) | (3.0, 2.0) | (6.0, 2.0) | (3.0, 0.5) | (3.0, 1.0) | (3.0, 3.0) |
|----------------------------------|------------|------------|------------|------------|------------|------------|
| $(F^*, K^*)$                     | (7, 9)     | (5, 7)     | (4, 6)     | (6, 8)     | (6, 8)     | (5, 7)     |
| $(\mu_B, \mu_V)$                 | (5.0, 3.0) | (5.0, 3.0) | (5.0, 3.0) | (5.0, 3.0) | (5.0, 3.0) | (5.0, 3.0) |
| $TC(F^*, K^*, \mu_B, \mu_V)$     | 830.281    | 794.920    | 781.249    | 875.276    | 828.551    | 780.993    |
| $\mu_B^*$                        | 6.070      | 5.889      | 5.800      | 5.222      | 5.592      | 5.947      |
| $\mu_V^*$                        | 3.970      | 4.132      | 4.321      | 7.757      | 6.454      | 0.933      |
| $TC(F^*, K^*, \mu_B^*, \mu_V^*)$ | 808.007    | 781.523    | 770.328    | 773.029    | 790.463    | 759.057    |



**Fig. 3.** Sensitivity analysis of TC with respect to  $\lambda$  for different  $F$  ( $\mu_B = 5.0, \mu_V = 3.0, \gamma = 3.0, \theta = 2.0$ ).



**Fig. 4.** Sensitivity analysis of TC with respect to  $\mu_B$  for different  $F$  ( $\lambda = 4.0, \mu_V = 3.0, \gamma = 3.0, \theta = 2.0$ ).

Case 7:  $\lambda = 4.0, \mu_B = 5.0, \mu_V = 3.0, \theta = 2.0, \gamma$  varies from 2.0 to 4.0.

Case 8:  $\lambda = 4.0, \mu_B = 5.0, \mu_V = 3.0, \gamma = 3.0, \theta$  varies from 1.0 to 3.0.

Figs. 3–7 show the sensitivity performance of the expected cost with respect to  $\lambda, \mu_B, \mu_V, \gamma$  and  $\theta$ , respectively. It is noted that the sign of sensitivity indicates an increase or decrease in the expect cost by changing the values of system parameters. In Fig. 3, it shows that (i)  $\partial TC / \partial \lambda$  is positive, which means that incremental change of  $\lambda$  increases the expected

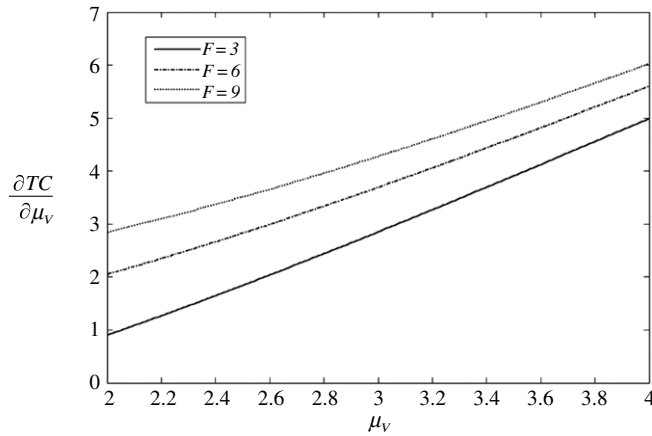


Fig. 5. Sensitivity analysis of TC with respect to  $\mu_V$  for different  $F$  ( $\lambda = 4.0, \mu_B = 5.0, \gamma = 3.0, \theta = 2.0$ ).

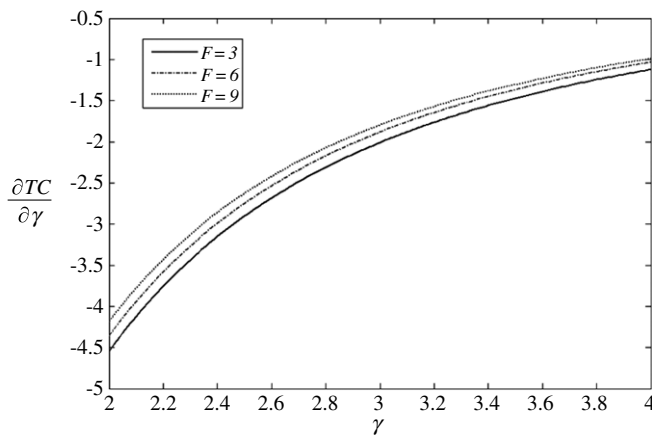


Fig. 6. Sensitivity analysis of TC with respect to  $\gamma$  for different  $F$  ( $\lambda = 4.0, \mu_B = 5.0, \mu_V = 3.0, \theta = 2.0$ ).

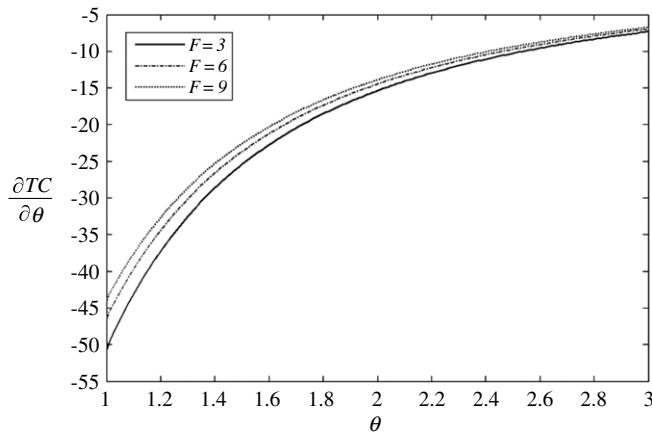


Fig. 7. Sensitivity analysis of TC with respect to  $\theta$  for different  $F$  ( $\lambda = 4.0, \mu_B = 5.0, \mu_V = 3.0, \gamma = 3.0$ ).

cost; (ii)  $\partial TC/\partial \lambda$  increases as  $\lambda$  increases for all  $F$ ; and (iii) as  $\lambda$  is fixed,  $\partial TC/\partial \lambda$  is getting larger as  $F$  decreases. We observe from Fig. 4 that (i)  $\partial TC/\partial \mu_B$  increases as  $\mu_B$  increases for all  $F$ ; and (ii)  $\partial TC/\partial \mu_B$  is negative when  $\mu_B$  is smaller than around 5.8 which means that an incremental change of  $\mu_B$  improves the expected cost. Fig. 5 shows that (i)  $\partial TC/\partial \mu_V$  is positive; (ii)  $\partial TC/\partial \mu_V$  increases as  $\mu_V$  increases; and (iii) as  $\mu_V$  is fixed,  $\partial TC/\partial \mu_V$  increases as  $F$  increases. It appears from Fig. 6 that (i)  $\partial TC/\partial \gamma$  is negative; (ii)  $\partial TC/\partial \gamma$  increases as  $\gamma$  increases for all  $F$ ; and (iii) as  $\gamma$  is fixed,  $\partial TC/\partial \gamma$  increases as  $F$  increases.

In Fig. 7, one can easily see that (i)  $\partial TC/\partial\theta$  is negative; (ii)  $\partial TC/\partial\theta$  increases as  $\theta$  increases; and (iii) as  $\theta$  is fixed,  $\partial TC/\partial\theta$  increases as  $F$  increases. In summary, it reveals from Figs. 3–7 that  $\lambda$  and  $\mu_B$  affect the expected cost significantly.

## 8. Conclusions

In this paper, we considered the  $F$ -policy M/M/1/K/WV queueing system with an exponential startup time. Steady-state probability vectors were developed in matrix form by the matrix-analytic method. We utilized the probability vectors to perform various system characteristics. An expected cost function per unit time was constructed to determine the optimal threshold value  $F^*$ , the optimal capacity  $K^*$ , the optimal service rates  $\mu_B^*$  and  $\mu_V^*$  at the minimum cost by using the direct method and Quasi-Newton method. Sensitivity analysis of the cost function has been done for specific values of the system parameters  $\lambda$ ,  $\mu_V$ ,  $\mu_B$ ,  $\theta$  and  $\gamma$ . Finally, numerical examples were presented to illustrate how to obtain the optimal solutions. Moreover, numerical investigations show that  $\lambda$  and  $\mu_B$  affect the expected cost significantly. Hence, these results could be helpful for the system analyst to make reliable decisions in reducing the cost.

## References

- [1] B.T. Doshi, Queueing system with vacations—a survey, *Queueing Systems* 1 (1986) 29–66.
- [2] H. Takagi, *Queueing Analysis: A Foundation of Performance Evaluation Vacation and Priority Systems, Part I*, North-Holland, Amsterdam, 1991.
- [3] L.D. Servi, S.G. Finn, M/M/1 queues with working vacations (M/M/1/WV), *Performance Evaluation* 50 (2002) 41–52.
- [4] S.M. Gupta, Interrelationship between controlling arrival and service in queueing systems, *Computers and Operations Research* 22 (1995) 1005–1014.
- [5] M. Yadin, P. Naor, Queueing systems with a removable service station, *Operational Research Quarterly* 14 (1963) 393–405.
- [6] D.P. Heyman, Optimal operating policies for M/G/1 queueing system, *Operations Research* 16 (1968) 362–382.
- [7] K.R. Balachandran, Control policies for a single server system, *Management Science* 19 (1973) 1013–1018.
- [8] F. Karaesmen, S.M. Gupta, Duality relations for queues with arrival and service control, *Computers and Operations Research* 24 (1997) 529–538.
- [9] K.-H. Wang, C.-C. Kuo, W.L. Pearn, Optimal control of an M/G/1/K queueing system with combined  $F$  Policy and startup time, *Journal of Optimization Theory and Applications* 135 (2007) 285–299.
- [10] K.-H. Wang, C.-C. Kuo, W.L. Pearn, A recursive method for the  $F$ -policy G/M/1/K queueing system with an exponential startup time, *Applied Mathematical Modelling* 32 (2008) 958–970.
- [11] J.D. Kim, D.W. Choi, K.C. Chae, Analysis of queue-length distribution of the M/G/1 queue with working vacations, in: *Hawaii International Conference on Statistics and Related Fields*, 2003.
- [12] J.-H. Li, N.-S. Tian, Z.-Y. Ma, Performance analysis of GI/M/1 queue with working vacations and vacation interruption, *Applied Mathematical Modelling* 31 (2007) 880–894.
- [13] X.W. Yi, J.D. Kim, D.W. Choi, K.C. Chae, The Geo/G/1 Queue with disasters and multiple working vacations, *Stochastic Models* 23 (2007) 537–549.
- [14] N. Tian, Z. Ma, M. Liu, The discrete time Geom/Geom/1 queue with multiple working vacations, *Applied Mathematical Modelling* 32 (2008) 2941–2953.
- [15] Y. Baba, Analysis of a GI/M/1 queue with multiple working vacations, *Operations Research Letters* 33 (2005) 201–209.
- [16] M. Jain, P.K. Agrawal, M/E<sub>k</sub>/1 queueing system with working vacation, *Quality Technology & Quantitative Management* 4 (2007) 455–470.
- [17] D. Wu, H. Takagi, M/G/1 queue with multiple working vacations, *Performance Evaluation* 63 (2006) 654–681.
- [18] A.D. Banik, U.C. Gupta, S.S. Pathak, On the GI/M/1/N queue with multiple working vacations—analytic analysis and computation, *Applied Mathematical Modelling* 31 (2007) 1701–1710.
- [19] J.-H. Li, N.-S. Tian, The discrete-time GI/Geo/1 queue with working vacations and vacation interruption, *Applied Mathematics and Computation* 185 (2007) 1–10.