



Optimal randomized control policy of an unreliable server system with second optional service and startup

Second optional
service and
startup

783

Dong-Yuh Yang

*Department of Industrial Engineering and Management,
National Chiao Tung University, Hsinchu, Taiwan*

Kuo-Hsiung Wang

*Department of Applied Mathematics,
National Chung-Hsing University, Taichung, Taiwan*

Jau-Chuan Ke

*Department of Applied Statistics,
National Taichung Institute of Technology,
Taichung, Taiwan, and*

Wen Lea Pearn

*Department of Industrial Engineering and Management,
National Chiao Tung University, Hsinchu, Taiwan*

Received 5 December 2007
Revised 22 February 2008
Accepted 3 March 2008

Abstract

Purpose – To study the optimization of a randomized control problem in an M/G/1 queue in which a removable and unreliable server may provide two phases of heterogeneous service to arriving customers.

Design/methodology/approach – Arriving customers follow a Poisson process and require the first essential service (FES). As soon as FES of a customer is completed, the customer may leave the system or opt for the second optional service (SOS). The service times of FES channel and SOS channel are assumed to be general distribution functions. The server requires a startup time with random length before starting service. When the server is working, he may meet unpredictable breakdowns but is immediately repaired. The inter-breakdown time and repair time of the removable server are exponentially random variable and generally random variable, respectively. By the convex combination property and the renewal reward theorem, several system performances are obtained. A cost model is developed to search the optimal two-threshold policy at a minimum cost. Sensitivity analysis is performed.

Findings – Expressions for various system performances are derived. Sensitivity analysis of optimal randomized control policy (based on the developed expected cost function) with respect to system parameters is investigated.

Originality/value – It is the first time that analytic results of sensitivity analysis of optimal randomized control policy for the complex system have been obtained which is quite useful and significant for engineers.

Keywords Sensitivity analysis, Optimal control

Paper type Research paper



Introduction

In this paper, we study the optimal control of a $\langle T, p \rangle$ -policy M/G/1 queue with second optional service (SOS) and general startup times in which the server is typically subject to unpredictable breakdowns. We say that the policy is a $\langle T, p \rangle$ -policy if it is prescribed by the following actions:

- switch the server off when the system becomes empty;
- if the server is turned off, he takes a vacation of time T whenever the system becomes empty. If at least one customer presents in the system switch the server on with probability p ($p \in [0, 1]$), and leave the server off with probability $(1-p)$. As the server is turned off, he will take another vacation of time T until the system becomes empty; and
- do not switch the server at other epochs.

In other words, the $\langle T, p \rangle$ -policy is to control the server randomly at the beginning epoch of the service when at least one customer appears.

By the definition of $\langle T, p \rangle$ -policy, a $\langle T, 1 \rangle$ -policy coincides with the T -policy introduced by Heyman (1977) and a $\langle T, 0 \rangle$ -policy coincides with the $2T$ -policy. One of the important characteristics in queueing systems is the service process. Madan (2000) introduced the ordinary M/G/1 queue with SOS that a customer may leave the system either after the first essential service (FES) with probability $(1-\eta)$ or at the completion of FES may immediately go for SOS with probability η ($0 < \eta < 1$). We will present this queue as the $\langle T, p \rangle$ -policy M/G(G,G)/1 queue, where the second symbol denotes service time distributions for both FES and SOS channels, the third symbol denotes the repair time distributions for both FES and SOS channels and the fourth symbol is the startup time distribution.

It is assumed that arriving customers follow a Poisson process with rate λ . Arriving customers form a single waiting line at a server based on the order of their arrivals; that is, in a first-come, first-served (FCFS) discipline. A single server need serve all arriving customers for FES. The service times of FES channel are independent and identically distributed (i.i.d.) random variables obeying a general distribution function $S_1(t)$ ($t \geq 0$) with a mean service time μ_{S_1} and a finite variance $\sigma_{S_1}^2$. As soon as FES of a customer is completed, a customer may leave the system with probability $1-\eta$ or may opt for SOS with probability η , at the completion of which the customer departs from the system and the next customer, if any, is taken up for his FES. The service times of SOS channel are another independent and identically distributed (i.i.d.) random variables having a general distribution function $S_2(t)$ ($t \geq 0$) with a mean service time μ_{S_2} and a finite variance $\sigma_{S_2}^2$. The server can serve only one customer at a time. When the server is working, it may meet unpredictable breakdowns but is immediately repaired. We assume that a server's breakdown time has an exponential distribution with rate α_1 in the FES channel. In the SOS channel, the server fails at an exponential rate α_2 . When the server fails, it is immediately repaired at a repair facility, where the repair times of FES and SOS channels are independent general distributions with distribution functions $R_1(t), R_2(t)$, ($t \geq 0$), and the mean repair times μ_{R_1}, μ_{R_2} , and the finite variances $\sigma_{R_1}^2, \sigma_{R_2}^2$, respectively. Service is allowed to be interrupted if the server breaks down, and the server is immediately repaired. Although no service occurs during the repair period of the server, customers continue to arrive following a Poisson process. Once the failed server is repaired, it immediately returns to serve a customer

until the system is empty. When the system is empty, the server operates $\langle T, p \rangle$ -policy. He requires a startup time with random length before starting service. Again, the startup times are independent and identically distributed random variables obeying a general distribution function $U(t)$ ($t \geq 0$) with a mean startup time μ_U and a finite variance σ_U^2 . Once the startup is terminated, the server begins serving the waiting customers until the system becomes empty.

Queueing systems with server vacations have received considerable attention in literatures since Levy and Yechiali (1975). Server vacations are useful for the system in which the server may utilize his idle time for additional tasks. There have been several contributions considering queueing systems with server vacations (Doshi, 1986; Takagi, 1991). The well-known T -policy with a reliable server for controlling the queueing system was first introduced by Heyman (1977) which is defined as follows: as soon as the system is empty, the server deactivates and leaves for a vacation with fixed length of time T . After a vacation period of length T , the server returns to the system. It begins to serve if there is at least one customer in the waiting line; otherwise, the server waits another period of length T and so on until at least one customer is present. Gakis *et al.* (1995) developed the distributions and the first two moments of the busy and idle periods in an M/G/1 queue operating under six dyadic policies. Tadj (2003) investigated an M/G/1 quorum queueing system under T -policy with a reliable server, where quorum is a bulk service that the server waits until the number of waiting customer reaches a fixed accumulation level r ($r \geq 1$). He derived the probability generating function of the number customers in the system and system characteristics. For an unreliable server, Gaver (1962) first proposed an ordinary M/G/1 queue with interrupted service and priorities. Sengupta (1990) extended Gaver's system to GI/G/1 case. Wang and Ke (2002) analyzed an M/G/1 queue with server breakdowns operating under the N -policy, T -policy and $\text{Min}(N, T)$ -policy. When operating a system, the server often requires a startup time before starting the service. The server startup corresponds to the pre-work of the server before starting the service. Doshi (1985) and Takagi (1991), respectively, studied GI/G/1 and M/G/1 queueing systems where the server requires a startup time before providing service. Later, Ke (2005) examined a modified T -policy for the M/G/1 queue with an unreliable server and startup where a single server may take most J vacations repeatedly until at least one customer appears in the queue upon returning from a vacation, and the server needs a startup time before starting each service period. Ke (2008a, b) extended Ke's model (2005) to $M^{[x]}/G/1$ queue with an unreliable server, startup and closedown. He presented an efficient solution algorithm to search the joint optimal two-threshold value so as to minimize the total expected cost function. The joint optimal thresholds of two NT vacation policies of an M/G/1 queue were examined by Ke (2006), in which server breakdown, startup and closedown are possibly considered and the length of the vacation period is controlled either by the number of arrivals during the vacation period, or by a timer.

For the randomized control policy, Feinberg and Kim (1996) investigated either $\langle p, N \rangle$ - or $\langle N, p \rangle$ -policy M/G/1 queue with a reliable sever at first. Kim and Moon (2006) considered the system with the (T, p) -policy, exploited its properties and found the optimal values of T and p for a constrained problem. In many real service systems, one encounters numerous examples of the queueing situations where all arrivals require the main service and only some may require the subsidiary service provided by the server. A pioneering work in this queueing situation is Madan (2000), who first

introduced the concept of *SOS*. Mandan (2000) examined the time-dependent as well as the steady-state behavior of an *M/G/1* queue with *SOS*, using the supplementary variable technique in which he considered general service time distribution for *FES* channel and exponential service time distribution for *SOS* channel. Medhi (2002) proposed an *M/G/1* queue with second optional channel and developed the explicit expressions for the mean queue length and the mean waiting time. Based on the supplementary variable technique, Wang (2004) studied the reliability behavior in an *M/G/1* queue with *SOS* and server breakdowns. Recently, Ke (2008a, b) analyzed an $M^{[x]}/G/1$ queue with a startup time and J additional options for service. He derived the explicit formulae for various system performances such as the expected number of customers in the system, the expected waiting time of an arbitrary customer in the queue, expected lengths of busy, idle periods and the expected unfinished work, etc.

The paper is organized as follows. In the next section, we develop the probability generating function and various system performances under T -policy *M/G/1* queue with *SOS*, server breakdowns and general startup times such as the expected number of customers in the system, the expected length of the idle, the startup, the busy and the breakdown periods, etc. In Section 3, we use the convex combination property and the renewal reward theorem to derive some exact and important system performances under the $\langle T, p \rangle$ -policy *M/G(G,G)/1* queue. In Section 4, we first construct the expected cost per unit time in the long run based on the derived system performances in the previous section. Then the optimal thresholds (T^* and p^*) of the $\langle T, p \rangle$ -policy *M/G(G,G)/1* queue can be analytically determined which minimize the expected cost per unit time. Sensitivity analysis and a numerical result are performed in Sections 5 and 6, respectively. Finally, Section 7 gives some concluding remarks.

System performances of the T -policy queue with *SOS*

In this section, we develop important system performances for T -policy *M/G(G,G)/1* queue, such as:

- the expected number of customers in the system;
- the expected length of the idle period;
- the expected length of the startup period and the completion period;
- the expected length of the busy period and the breakdown period; and
- the expected length of the busy cycle.

Expected number of customers in the system

Let us excerpt some important results of Medhi (2002) as follows:

$$E[S] = \mu_{S_1} + \eta\mu_{S_2}, \tag{1}$$

$$E[S^2] = E[S_1^2] + 2\eta\mu_{S_1}\mu_{S_2} + \eta E[S_2^2], \tag{2}$$

where S is the (total) service time random variable. $E[S_1^2]$ and $E[S_2^2]$, respectively, denote the second moment of the service time distributions of *FES* and *SOS* channels.

Let H_1 and H_2 be a random variable representing the completion time of *FES* and *SOS* channels. The completion time of a customer consists of the service time of a customer and the repair time of a server. Using the known results of Wang and Ke's

(2002) paper, we get the first two moments of the completion time distribution for FES and SOS channels:

Second optional
service and
startup

$$E[H_i] = \mu_{S_i}(1 + \alpha_i \mu_{R_i}), \quad i = 1, 2, \quad (3)$$

$$E[H_i^2] = (1 + \alpha_i \mu_{R_i})^2 E[S_i^2] + \alpha_i \mu_{S_i} E[R_i^2], \quad i = 1, 2, \quad (4)$$

787

where $E[R_1^2]$ and $E[R_2^2]$, respectively, denote the second moment of the repair time distributions of FES and SOS channels.

Applying the Medhi's formula (2002) for the probability generating function (p.g.f.) of the number of customers in an M/G/1 queue with SOS and reliable channels, we obtain the p.g.f. of the number of customers in an M/G/1 queue with SOS and server breakdowns as follows:

$$G(z) = \frac{(1 - \rho_H)(1 - z)H^*(\lambda - \lambda z)}{H^*(\lambda - \lambda z) - z}, \quad (5)$$

where H is a random variable denoting the (total) completion time of a customer, $\rho_H = \lambda E[H] = \lambda(E[H_1] + \eta E[H_2])$ and $H^*(\cdot)$ is the Laplace-Stieltjes transform (abbreviated LST) of the total completion time for a customer.

Let Y_t be the number of customers who arrive at the system during $[0, t]$. Using the definition of Poisson arrivals, $e^{-\lambda t}(\lambda t)^n/n!$ is the probability that n customers arrive during $[0, t]$. Let S_n denote the epoch of the n th arriving customer and $F_n(t)$ be its distribution function. Then we get:

$$F_n(t) = \int_0^t \frac{\lambda(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda x} dx = 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t} = \Pr\{Y_t \geq n\}. \quad (6)$$

It is obviously that:

$$\Pr\{Y_t = n\} = F_n(t) - F_{n+1}(t), \quad n = 0, 1, 2, \dots \quad (7)$$

Given a period of length T , the distribution of S_n can be represented as:

$$F_n(T) = \Pr\{Y_T \geq n\} = 1 - \sum_{k=0}^{n-1} \frac{(\lambda T)^k}{k!} e^{-\lambda T}. \quad (8)$$

Let $G_B(z)$ be the p.g.f. of the number of customers who arrive during a period T . Thus, one may write:

$$G_B(z) = \sum_{n=0}^{\infty} z^n \Pr\{Y_T = n\} = e^{-(1-z)\lambda T}. \quad (9)$$

After a period length T , the server begins to perform a startup once a customer arrives. Let ϕ denote the number of customers who arrive during a period length T and startup period. The probability distribution is:

$$\varphi_k = \Pr(\phi = k), \quad k = 0, 1, 2, \dots \quad (10)$$

Since the Poisson process from any point on is independent of all that has previously occurred, we use $\phi(z)$ to denote the p.g.f. of ϕ , which can be expressed as:

$$\varphi(z) = \sum_{k=0}^{\infty} \varphi_k z^k = G_B(z)U^*(\lambda - \lambda z) = e^{-(1-z)\lambda T}U^*(\lambda - \lambda z), \quad (11)$$

where $U^*(\cdot)$ is the LST of U .

According to the well-known stochastic decomposition results by Fuhrmann and Cooper (1985) and the above inferences listed, we obtain the p.g.f. of the number of customers found in the T -policy M/G(G,G)/1 queue as follows:

$$G_T(z) = G(z) \left[\frac{1 - \varphi(z)}{\varphi'(1)(1 - z)} \right], \quad (12)$$

where $G(z)$ is given in Equation (5). Let L_T be the expected number of customers in the T -policy M/G(G,G)/1 queue. Thus, it follows that:

$$L_T = G'_T(z)|_{z=1} = \frac{1}{(T + \mu_U)} \left[\frac{\lambda T^2}{2} + \rho_U T + \frac{\lambda E(U^2)}{2} \right] + L_H, \quad (13)$$

where $\rho_U = \lambda \mu_U$ and:

$$L_H = \rho_H + \frac{\lambda^2 E(H^2)}{2(1 - \rho_H)}. \quad (14)$$

It has to be noted that ρ_H is assumed to be less than unity.

Expected length of the idle period

The idle period begins when all the customers in the system are served and no customers are waiting for service. It terminates when at least one customer arrives at the period length T . We can easily see that:

$$E[I_T] = T. \quad (15)$$

Expected length of the startup and completion periods

The server begins startup when there is at least one waiting customer at the end of the fixed period T in the system. We call this startup period and denote it by U_T . It follows that:

$$E[U_T] = \mu_U. \quad (16)$$

The completion period is from the end of the startup period to the epoch of no customers in the system, which occurs before the system becomes empty and can be represented as the sum of the busy period and the breakdown period. A time interval when the server is working continuously is called a busy period. During the busy period, the server may break down when FES or SOS is provided and start its repair immediately. This is called the breakdown period. After the server is repaired, it returns and provides service until there are no customers in the system.

First let $H^*_O(\theta)$ denote the LST of the completion period in the ordinary M/G/1 queue with SOS. Applying an important result of Kleinrock (1975), Tang (1997) and Wang (1997) demonstrated that the LST of the completion period started with one customer in the M/G/1 queue with SOS can be represented as:

$$H_O^*(\theta) = H^*[\theta + \lambda - \lambda H_O^*(\theta)]. \tag{17}$$

Differentiating (17) with respect to θ , we derive the expected length of the completion period for the ordinary M/G/1 queue with SOS as:

$$E[H_O] = -\left. \frac{dH_O^*(\theta)}{d\theta} \right|_{\theta=0} = \frac{E[H]}{1 - \lambda E[H]} = \frac{\mu_{S_1}(1 + \alpha_1 \mu_{R_1}) + \eta \mu_{S_2}(1 + \alpha_2 \mu_{R_2})}{1 - \rho_H}. \tag{18}$$

Then let $H_T^*(\theta)$ be the LST of the completion period for the T -policy M/G(G,G)/1 queue. It follows from a property of Poisson arrival process and the assumption of exhaustive service that the instant of commencement of each idle period is the regeneration point of the system. If there are k customers in the system at the end of the startup period, the successive completion period will consist of k independent completion periods. Thus, we have:

$$H_T^*(\theta) = \sum_{k=0}^{\infty} \varphi_k [H_O^*(\theta)]^k = \varphi [H_O^*(\theta)], \tag{19}$$

Substituting Equation (11) into Equation (19), we obtain:

$$H_T^*(\theta) = e^{-[1-H_O^*(\theta)]\lambda T} U^*(\lambda - \lambda H_O^*(\theta)). \tag{20}$$

The expected length of the completion period is given by:

$$E[H_T] = \frac{(T + \mu_U)\rho_H}{1 - \rho_H} = \frac{(\lambda T + \rho_U)[\mu_{S_1}(1 + \alpha_1 \mu_{R_1}) + \eta \mu_{S_2}(1 + \alpha_2 \mu_{R_2})]}{1 - \lambda E[H]}. \tag{21}$$

Expected length of the busy and breakdown periods

We denote the expected length of the busy and breakdown periods by $E[B_T]$ and $E[D_T]$, respectively. Recall that the completion period is composed of the busy period and the breakdown period which implies that $E[H_T] = E[B_T] + E[D_T]$. Hence from (21), we obtain:

$$E[B_T] = \frac{(\lambda T + \rho_U)(\mu_{S_1} + \eta \mu_{S_2})}{1 - \lambda E[H]}, \tag{22}$$

and:

$$E[D_T] = \frac{(\lambda T + \rho_U)(\alpha_1 \mu_{S_1} \mu_{R_1} + \eta \alpha_2 \mu_{S_2} \mu_{R_2})}{1 - \lambda E[H]}. \tag{23}$$

Expected length of the busy cycle

The expected length of busy cycle for the T -policy M/G(G,G)/1 queue is denoted by $E[C_T]$. Since the busy cycle consists of the idle period ($E[I_T]$), the startup period ($E[U_T]$), the busy period ($E[B_T]$) and the breakdown period ($E[D_T]$). Hence, it can be shown that:

$$E[C_T] = E[I_T] + E[U_T] + E[B_T] + E[D_T] = \frac{T + \mu_U}{1 - \rho_H}. \quad (24)$$

The $\langle T, p \rangle$ -policy M/G (G, G)/1 Queue

The primary objective of this section is to develop the various system performances for the $\langle T, p \rangle$ -policy M/G (G, G)/1 queue, including:

- the expected length of the idle, startup, busy, breakdown periods and busy cycle;
- the long-run fraction of time measures; and
- the expected number of customers in the system.

Expected length of the idle, startup, busy, breakdown periods and busy cycle

We denote by $(I_{2T}, U_{2T}, B_{2T}, D_{2T})$ and $(I_{T,p}, U_{T,p}, B_{T,p}, D_{T,p})$ the idle, startup, busy, breakdown periods for the $2T$ -policy and $\langle T, p \rangle$ -policy M/G(G,G)/1 queue, respectively. Let C_{2T} and $C_{T,p}$ be the busy cycle for the $2T$ -policy and $\langle T, p \rangle$ -policy M/G(G,G)/1 queues, respectively. It follows from Feinberg and Kim (1996) that the system performances for the $\langle T, p \rangle$ -policy queue are the convex combinations of the system performances for the T -policy queue and the $2T$ -policy queue. Using the above formulas (15), (16) and (22)-(24), we have:

$$E[I_{T,p}] = pE[I_T] + (1 - p)E[I_{2T}] = T(2 - p), \quad (25)$$

$$E[U_{T,p}] = pE[U_T] + (1 - p)E[U_{2T}] = \mu_U, \quad (26)$$

$$E[B_{T,p}] = pE[B_T] + (1 - p)E[B_{2T}] = \frac{[\lambda T(2 - p) + \rho_U](\mu_{S_1} + \eta\mu_{S_2})}{1 - \rho_H}, \quad (27)$$

$$\begin{aligned} E[D_{T,p}] &= pE[D_T] + (1 - p)E[D_{2T}] \\ &= \frac{[\lambda T(2 - p) + \rho_U](\alpha_1\mu_{S_1}\mu_{R_1} + \eta\alpha_2\mu_{S_2}\mu_{R_2})}{1 - \rho_H}, \end{aligned} \quad (28)$$

$$E[C_{T,p}] = pE[C_T] + (1 - p)E[C_{2T}] = \frac{T(2 - p) + \mu_U}{1 - \rho_H}. \quad (29)$$

Thus, we have the number of busy cycle per unit time:

$$\frac{1}{E[C_{T,p}]} = \frac{1 - \rho_H}{T(2 - p) + \mu_U}. \quad (30)$$

The long-run fraction of time measures

In steady-steady, let:

- $P_{I_{TP}}$ \equiv the probability that the server is idle (turned-off).
- $P_{U_{TP}}$ \equiv the probability that the server is startup.
- $P_{B_{TP}}$ \equiv the probability that the server is busy.
- $P_{D_{TP}}$ \equiv the probability that the server is broken down.

From Equations (25)-(29), we get:

$$P_{I_{TP}} = \frac{E[I_{T,\phi}]}{E[C_{T,\phi}]} = \frac{T(1 - \rho_H)(2 - \phi)}{T(2 - \phi) + \mu_U}, \quad (31)$$

$$P_{U_{TP}} = \frac{E[U_{T,\phi}]}{E[C_{T,\phi}]} = \frac{\mu_U(1 - \rho_H)}{T(2 - \phi) + \mu_U}, \quad (32)$$

$$P_{B_{TP}} = \frac{E[B_{T,\phi}]}{E[C_{T,\phi}]} = \lambda(\mu_{S_1} + \eta\mu_{S_2}), \quad (33)$$

$$P_{D_{TP}} = \frac{E[D_{T,\phi}]}{E[C_{T,\phi}]} = \lambda(\alpha_1\mu_{S_1}\mu_{R_1} + \eta\alpha_2\mu_{S_2}\mu_{R_2}). \quad (34)$$

Expected number of customers in the system

We denote Π_T , Π_{2T} and $\Pi_{T,\phi}$ by the cumulative amount of time that all customers spent in the system during a busy cycle for the T -, $2T$ - and $\langle T, \phi \rangle$ -policy M/G(G,G)/1 queue, respectively. From the renewal reward theorem, we obtain:

$$E[\Pi_T] = L_T E[C_T] = \frac{1}{1 - \rho_H} \left[\frac{\lambda T^2}{2} + T\rho_U + \frac{\lambda E(U^2)}{2} + L_H(T + \mu_U) \right], \quad (35)$$

where L_H is given in (14). It follows that:

$$\begin{aligned} E[\Pi_{T,\phi}] &= \phi E[\Pi_T] + (1 - \phi)E[\Pi_{2T}] \\ &= \frac{1}{1 - \rho_H} \left[\lambda T^2 \left(2 - \frac{3}{2}\phi \right) + T\rho_U(2 - \phi) \right. \\ &\quad \left. + \frac{\lambda E(U^2)}{2} + L_H[T(2 - \phi) + \mu_U] \right]. \end{aligned} \quad (36)$$

Let $L_{T,\phi}$ denote the expected number of customers in the $\langle T, \phi \rangle$ -policy M/G(G,G)/1 queue. Again, from the renewal reward theorem, we have:

$$\begin{aligned} L_{T,\phi} &= \frac{E[\Pi_{T,\phi}]}{E[C_{T,\phi}]} \\ &= \frac{1}{T(2 - \phi) + \mu_U} \left[\lambda T^2 \left(2 - \frac{3}{2}\phi \right) + T\rho_U(2 - \phi) + \frac{\lambda E(U^2)}{2} \right] + L_H. \end{aligned} \quad (37)$$

It follows from Feinberg and Kim (1996) that $L_{T,\phi}$ is a convex combination of L_T for a T -policy and L_{2T} for a $2T$ -policy. Thus, we get:

$$L_{T,\phi} = \Theta L_T + (1 - \Theta)L_{2T}, \quad (38)$$

where:

$$\Theta = \frac{\phi(T + \mu_U)}{(2 - \phi)T + \mu_U}.$$

After some algebraic manipulation, it is significant to see that expression (38) is in agreement with expression (37), which confirms the results by Feinberg and Kim (1996).

From the explicit formula derived in (37), we will demonstrate the following corollary. First, differentiating $L_{T,p}$ with respect to p , we obtain:

$$\frac{\partial L_{T,p}}{\partial p} = \frac{-\lambda T [2T^2 + 3T\mu_U + (\mu_U^2 - \sigma_U^2)]}{2[T(2-p) + \mu_U]^2} \quad \text{in } p \in [0, 1], \quad (39)$$

where $\sigma_U^2 = E(U^2) - \mu_U^2$. For $T > (-3\mu_U + \sqrt{\mu_U^2 + 8\sigma_U^2})/4$, we can prove that $(\partial L_{T,p}/\partial p) < 0$. Thus, $L_{T,p}$ is strictly decreasing in $p \in [0, 1]$ if $T > (-3\mu_U + \sqrt{\mu_U^2 + 8\sigma_U^2})/4$.

Next, differentiating $L_{T,p}$ with respect to T , yields:

$$\frac{\partial L_{T,p}}{\partial T} = \frac{\lambda [(2-p)(2-\frac{3}{2}p)T^2 + (4-3p)\mu_U T + (2-p)(\mu_U^2 - \sigma_U^2)/2]}{[T(2-p) + \mu_U]^2}. \quad (40)$$

For $p \in [0, 1]$, we demonstrate that $L_{T,p}$ is strictly increasing in T when T is fulfilled the following condition: $T > [-\mu_U + \sqrt{\mu_U^2 - (2-p)^2(\mu_U^2 - \sigma_U^2)/(4-3p)}]/(2-p)$. Finally, for $p \in [0, 1]$ and $q \in [0, 1]$, we obtain from Equation (37):

$$\begin{aligned} L_{2T,q} - L_{T,p} = & \left\{ \lambda T^3 [2(4-3q)(2-p) - (4-3p)(2-q)] + \lambda T^2 \mu_U \right. \\ & \times \left(6 - 6q + \frac{3}{2}p \right) + (\lambda T/2) [2(1-q) + p] [\mu_U^2 - \sigma_U^2] \left. \right\} \\ & \times \{ [2T(2-q) + \mu_U] [T(2-p) + \mu_U] \}^{-1}. \end{aligned} \quad (41)$$

Substituting $p = 0$ and $q = 1$ into Equation (41), it follows that $L_{2T,q} = L_{T,p}$. On the contrary, setting $L_{2T,q} = L_{T,p}$, we can derive $p = 0$ and $q = 1$. From the listed above results, we have the corollary as follows:

Corollary 1:

- For $T > (-3\mu_U + \sqrt{\mu_U^2 + 8\sigma_U^2})/4$, $L_{T,p}$ is a strictly decreasing function in $p \in [0, 1]$.
- For $p \in [0, 1]$, $L_{T,p}$ is a strictly increasing function in $T > [-\mu_U + \sqrt{\mu_U^2 - (2-p)^2(\mu_U^2 - \sigma_U^2)/(4-3p)}]/(2-p)$.
- $L_{T,p} = L_{2T,q}$ if and only if $p = 0$ and $q = 1$.

Optimal $\langle T, p \rangle$ -policy

We develop the expected cost function per unit time for the $\langle T, p \rangle$ -policy M/G(G,G)/1 queue, in which p and T are decision variables. Our objective is to determine the join optimal threshold values (say (T^*, p^*)), so as to minimize this function. Let us define the cost elements as follows:

- C_h , holding cost per unit time for each customer present in the system.
- C_f , cost per unit time for keeping the server off

- C_S , startup cost per unit time for the preparatory work of the server before starting the service.
- C_b , setup cost per busy cycle.

Without loss of generality, we assume that $C_S > C_f$. Utilizing the definition of each cost element listed above and its corresponding system performances, the expected cost with threshold (T, p) is given by:

$$F_0(T, p) = C_h L_{T,p} + C_f P_{I_{TP}} + C_s P_{U_{TP}} + C_l \frac{1}{E(C_{T,p})},$$

where $L_{T,p}$ is given in Equation (37). Since L_H is independent of T and p , we omit this term. Now, we are interested in obtaining the joint optimal thresholds (T^*, p^*) and to minimize $F_0(T, p)$ is equivalent to minimize the following equation:

$$\begin{aligned} F(T, p) &= \frac{1}{T(2-p) + \mu_U} \left\{ C_h \left[\lambda T^2 \left(2 - \frac{3}{2} p \right) + T \rho_U (2-p) + \lambda E(U^2)/2 \right] \right. \\ &\quad \left. + [C_f T(2-p) + C_S \mu_U + C_l](1 - \rho_H) \right\} \\ &= [\lambda \mu_U C_h + C_f(1 - \rho_H)] \\ &\quad + \frac{1}{2} \lambda C_h \left\{ 3T - \frac{1}{T(2-p) + \mu_U} [2T^2 + 3T\mu_U + \delta] \right\}, \end{aligned} \quad (42)$$

where $\delta = \mu_U^2 - \sigma_U^2 - 2[(C_S - C_f)\mu_U + C_l](1 - \rho_H)/\lambda C_h$.

In order to find the joint optimal thresholds (T^*, p^*) , we further define:

$$\begin{aligned} M(T, p) &= 3T - \frac{1}{T(2-p) + \mu_U} (2T^2 + 3T\mu_U + \delta) \\ &= 3T - \frac{1}{T(2-p) + \mu_U} (T - \omega_1)(T - \omega_2), \end{aligned} \quad (43)$$

where $\omega_1 = (-3\mu_U - \sqrt{9\mu_U^2 - 8\delta})/4$ and $\omega_2 = (-3\mu_U + \sqrt{9\mu_U^2 - 8\delta})/4$.

Obviously, the following results can be derived:

- As $0 < T < \omega_2$, $M(T, p)$ is an increasing function in $p \in [0, 1]$. It implies that:

$$\begin{aligned} h(T) &= \min_{0 \leq p \leq 1} M(T, p) \\ &= M(T, 0) \\ &= 3T - \frac{1}{2T + \mu_U} (2T^2 + 3T\mu_U + \delta), \quad 0 < T < \omega_2. \end{aligned} \quad (44)$$

- As $T = \omega_2 > 0$, it yields that $M(T, p) = 3T = 3\omega_2$ in $p \in [0, 1]$.
- As $T > \omega_2$, $M(T, p)$ is a decreasing function in $p \in [0, 1]$. We have:

$$g(T) = \min_{0 \leq p \leq 1} M(T, p) = M(T, 1) = 3T - \frac{1}{T + \mu_U} (2T^2 + 3T\mu_U + \delta), \quad T > \omega_2. \quad (45)$$

• Differentiating $h(T)$ in (44) with respect to T , we get:

$$\frac{dh(T)}{dT} = 3 - \frac{1}{(2T + \mu_U)^2} (4T^2 + 4T\mu_U + 3\mu_U^2 - 2\delta). \quad (46)$$

Setting $dh(T)/dT = 0$ yields $T_1^* = (-\mu_U + \sqrt{\mu_U^2 - \delta})/2$. Since:

$$\left. \frac{d^2h(T)}{dT^2} \right|_{T=T_1^*} = \frac{8\sigma_U^2 + (16[(C_S - C_f)\mu_U + C_1](1 - \rho_H)/\lambda C_h)}{(\sqrt{\mu_U^2 - \delta})^3} > 0, \quad (47)$$

$h(T)$ is a concave upward (convex) function. Thus, T_1^* is the unique minimum of $h(T)$ and the minimum of $h(T)$ can be expressed as:

$$h(T_1^*) = -2\mu_U + 2\sqrt{\sigma_U^2 + \frac{2[(C_S - C_f)\mu_U + C_1](1 - \rho_H)}{\lambda C_h}}, \quad 0 < T < \omega_2. \quad (48)$$

• Differentiating $g(T)$ in (45) with respect to T , it follows that:

$$\frac{dg(T)}{dT} = 3 - \frac{1}{(T + \mu_U)^2} (2T^2 + 4T\mu_U + 3\mu_U^2 - \delta). \quad (49)$$

Setting $dg(T)/dT = 0$ yields $T_2^* = -\mu_U + \sqrt{\mu_U^2 - \delta}$. Since:

$$\left. \frac{d^2g(T)}{dT^2} \right|_{T=T_1^*} = \frac{2\sigma_U^2 + (4[(C_S - C_f)\mu_U + C_1](1 - \rho_H)/\lambda C_h)}{(\sqrt{\mu_U^2 - \delta})^3} > 0, \quad (50)$$

$g(T)$ is also a concave upward (convex) function. Therefore, T_2^* is the unique minimum of $g(T)$ and the minimum of $g(T)$ is given by:

$$g(T_2^*) = -2\mu_U + 2\sqrt{\sigma_U^2 + \frac{2[(C_S - C_f)\mu_U + C_1](1 - \rho_H)}{\lambda C_h}}, \quad T > \omega_2. \quad (51)$$

Subsequently based upon (42)-(51), we can obtain the minimal value of $F(T, p)$, say F^* , which can be written as:

$$\begin{aligned} F^* &= [\lambda\mu_U C_h + C_f(1 - \rho_H)] + \frac{1}{2} \lambda C_h h(T_1^*) \\ &= [\lambda\mu_U C_h + C_f(1 - \rho_H)] + \frac{1}{2} \lambda C_h g(T_2^*) \\ &= [\lambda\mu_U C_h + C_f(1 - \rho_H)] \\ &\quad + \frac{1}{2} \lambda C_h \left[-2\mu_U + 2\sqrt{\sigma_U^2 + \frac{2[(C_S - C_f)\mu_U + C_1](1 - \rho_H)}{\lambda C_h}} \right] \\ &= F(T_1^*, 0) = F(T_2^*, 1). \end{aligned} \quad (52)$$

Summarizing the listed above results, the joint optimal threshold values (T^*, p^*) that minimize (42) can be obtained by the following theorem:

Theorem 1. Let (T^*, p^*) be the joint optimal threshold values that minimize the expected cost in (42), i.e. $F(T^*, p^*)$ is minimum. If we define Ψ by:

$$\Psi = \begin{cases} 1/2, & \text{if } p = 0, \\ 1, & \text{if } p = 1. \end{cases}$$

Then p^* is equal to 0 or 1, and T^* is equal to $\Psi(-\mu_U + \sqrt{\mu_U^2 - \delta})$. That is:

$$T^* = \Psi \left(-\mu_U + \sqrt{\sigma_U^2 + \frac{2[(C_S - C_f)\mu_U + C_1](1 - \rho_H)}{\lambda C_h}} \right). \quad (53)$$

Sensitivity analysis

A system analyst often concerns with how the system performances can be affected by the changes of the input parameters. Consequently, a major advantage of sensitivity analysis on the queueing system is that they provide some guidance for the system analyst. We now perform a sensitivity analysis for changes in the optimal values T^* along with changes in specific values of the system parameters $\lambda, \mu_1, \mu_2, \alpha_1, \alpha_2, \beta_1, \beta_2, \eta, \gamma$ and the cost parameters C_h, C_f, C_s, C_i , where $\mu_{S_1} = 1/\mu_1, \mu_{S_2} = 1/\mu_2, \mu_{R_1} = 1/\beta_1, \mu_{R_2} = 1/\beta_2$ and $\mu_U = 1/\gamma$.

Sensitivity for λ

From Equation (53), differentiating T^* with respect to λ , we obtain:

$$\frac{\partial T^*}{\partial \lambda} = \frac{-\Psi\Theta}{\lambda^2 \sqrt{\sigma_U^2 + \frac{2\Theta(1 - \rho_H)}{\lambda}}} < 0, \quad (54)$$

where $\Theta = [(C_S - C_f)\mu_U + C_1]/C_h$. It implies that T^* decreases in λ .

Sensitivity for μ_1 and μ_2

Differentiating T^* with respect to μ_1 and μ_2 , respectively, yields:

$$\frac{\partial T^*}{\partial \mu_1} = \frac{\Psi\Theta(1 + \alpha_1\mu_{R_1})}{\mu_1^2 \sqrt{\sigma_U^2 + (2\Theta(1 - \rho_H)/\lambda)}} > 0, \quad (55)$$

$$\frac{\partial T^*}{\partial \mu_2} = \frac{\eta\Psi\Theta(1 + \alpha_2\mu_{R_2})}{\mu_2^2 \sqrt{\sigma_U^2 + (2\Theta(1 - \rho_H)/\lambda)}} > 0. \quad (56)$$

This means that T^* increases in μ_1 and μ_2 from Equations (55) and (56), respectively.

Sensitivity for α_1 and α_2

Differentiating T^* with respect to α_1 and α_2 , respectively, we get:

$$\frac{\partial T^*}{\partial \alpha_1} = \frac{-\Psi\Theta\mu_{S_1}\mu_{R_1}}{\sqrt{\sigma_U^2 + (2\Theta(1 - \rho_H)/\lambda)}} < 0, \tag{57}$$

$$\frac{\partial T^*}{\partial \alpha_2} = \frac{-\eta\Psi\Theta\mu_{S_2}\mu_{R_2}}{\sqrt{\sigma_U^2 + (2\Theta(1 - \rho_H)/\lambda)}} < 0. \tag{58}$$

The above results show that T^* decreases in α_1 and α_2 , respectively.

Sensitivity for β_1 and β_2

Differentiating T^* with respect to β_1 and β_1 , respectively, it follows that:

$$\frac{\partial T^*}{\partial \beta_1} = \frac{\Psi\Theta\alpha_1\mu_{R_1}}{\beta_1^2\sqrt{\sigma_U^2 + (2\Theta(1 - \rho_H)/\lambda)}} > 0, \tag{59}$$

$$\frac{\partial T^*}{\partial \beta_2} = \frac{\eta\Psi\Theta\alpha_2\mu_{R_2}}{\beta_2^2\sqrt{\sigma_U^2 + (2\Theta(1 - \rho_H)/\lambda)}} > 0. \tag{60}$$

From Equations (59) and (60), T^* increases in β_1 and β_1 , respectively.

Sensitivity for η

Differentiating T^* with respect to η is given by:

$$\frac{\partial T^*}{\partial \eta} = \frac{-\Psi\Theta\mu_{S_2}(1 + \alpha_2\mu_{R_2})}{\sqrt{\sigma_U^2 + (2\Theta(1 - \rho_H)/\lambda)}} < 0. \tag{61}$$

In Equation (61), we observe that T^* is decreasing in η .

Sensitivity for γ

Without loss of generality, it is assumed that a startup time is a positive random variable. An important quantity of a startup time is the coefficient of variation, which is a (dimensionless) measure of the variability of a startup time, defined as $\xi = \sigma_U/\mu_U$. In (53), $\xi\mu_U$ can be substituted for σ_U . Differentiating T^* with respect to μ_U , we have:

$$\frac{\partial T^*}{\partial \mu_U} = \Psi \left[-1 + \frac{2\xi^2\mu_U + \Delta}{2\sqrt{\xi^2\mu_U^2 + \Delta\mu_U + \Sigma}} \right], \tag{62}$$

where $\Delta = 2(C_S - C_i)(1 - \rho_H)/(\lambda C_h)$ and $\Sigma = 2C_1(1 - \rho_H)/\lambda C_h$. Let $\Gamma(\mu_U)$ be a function of μ_U , and it is expressed as $\Gamma(\mu_U) = (2\xi^2\mu_U + \Delta)^2 - 4(\xi^2\mu_U^2 + \Delta\mu_U)$. Recall that μ_U is equal to $1/\gamma$. From (62), three different conditions of $\Gamma(\mu_U)$ lead to different results of $\partial T^*/\partial \mu_U$ as follows:

- If $\Gamma(\mu_U) > 4\Sigma$, we have $\partial T^*/\partial \mu_U > 0$. It shows that T^* is increasing in μ_U (or decreasing in γ).
- If $\Gamma(\mu_U) = 4\Sigma$, it indicates that $\partial T^*/\partial \mu_U = 0$. It follows that T^* is independent of μ_U (or γ).
- If $\Gamma(\mu_U) < 4\Sigma$, it reveals that $\partial T^*/\partial \mu_U < 0$. It can be seen that T^* is decreasing in μ_U (or increasing in γ).

The relationship between $\Gamma(\mu_U)$ and 4Σ is applied to illustrate the sensitivity of μ_U (or γ). We provide four examples for four different startup time distributions such as deterministic, 2-stage Erlang, exponential and 2-stage hyper-exponential, respectively.

Example 1. (Deterministic). When a startup time obeys a deterministic distribution, it follows that $\xi = 0$ and $\Gamma(\mu_U) = \Delta^2 - 4\Delta\mu_U$. From discussions above, we have:

- T^* is increasing in μ_U (or decreasing in γ) when $\mu_U < \Delta/4 - C_1/(C_s - C_f)$.
- T^* is independent of μ_U (or γ) when $\mu_U = \Delta/4 - C_1/(C_s - C_f)$.
- T^* is decreasing in μ_U (or increasing in γ) when $\mu_U > \Delta/4 - C_1/(C_s - C_f)$.

Example 2. (2-stage Erlang). If the distribution of a startup time is a 2-stage Erlang distribution, ξ is equal to $1/\sqrt{2}$. Thus, $\Gamma(\mu_U)$ can be written as $-\mu_U^2 - 2\mu_U\Delta + \Delta^2$. After some algebraic manipulation, it implies that:

- T^* is increasing in μ_U (or decreasing in γ) when $\mu_U < -\Delta + \sqrt{2\Delta^2 - 4\Sigma}$.
- T^* is independent of μ_U (or γ) when $\mu_U = -\Delta + \sqrt{2\Delta^2 - 4\Sigma}$.
- T^* is decreasing in μ_U (or increasing in γ) when $\mu_U > -\Delta + \sqrt{2\Delta^2 - 4\Sigma}$.

Example 3. (Exponential). As a startup time follows an exponential distribution, we have $\xi = 1$ and $\Gamma(\mu_U) = \Delta^2$. It results in:

- T^* is increasing in μ_U (or decreasing in γ) when $\Delta > 4C_1/(C_s - C_f)$.
- T^* is independent of μ_U (or γ) when $\Delta = 4C_1/(C_s - C_f)$.
- T^* is decreasing in μ_U (or increasing in γ) when $\Delta < 4C_1/(C_s - C_f)$.

Example 4. (2-stage hyper-exponential). While a startup time belongs to a k -stage hyper-exponential distribution, ξ is greater than or equal to 1. Taking a 2-stage hyper-exponential distribution for example, we suppose that the mean is μ in stage 1 selected with probability 0.235 and 10μ in stage 2 selected with probability 0.765. This distribution has a coefficient of variation 2 ($\xi = 2$). $\Gamma(\mu_U)$ can be calculated as $48\mu_U^2 + 12\Delta\mu_U + \Delta^2$. It follows that:

- T^* is increasing in μ_U (or decreasing in γ) when $\mu_U > [-3\Delta + \sqrt{48\Sigma - 3\Delta^2}]/24$.
- T^* is independent of μ_U (or γ) when $\mu_U = [-3\Delta + \sqrt{48\Sigma - 3\Delta^2}]/24$.
- T^* is decreasing in μ_U (or increasing in γ) when $\mu_U < [-3\Delta + \sqrt{48\Sigma - 3\Delta^2}]/24$.

On the other hand, we can easily see from Equation (53) that T^* is increasing in C_i , C_s and decreasing in C_k , C_f . To close this chapter, we summarize the analytic sensitivity analysis in the following:

- T^* decreases in λ .
- T^* increases in μ_1 and μ_2 .

- T^* decreases in α_1 and α_1 .
- T^* increases in β_1 and β_1 .
- T^* decreases in η .
- T^* varies in γ by the relationship between $\Gamma(\mu_U)$ and 4Σ .
- T^* increases in C_1 and C_s , whereas decreases in C_k and C_f .

Numerical result

We will now give a numerical result in this section to demonstrate how to make decisions regarding the joint threshold values (T^*, p^*) , which minimize the expected cost function per unit time. Considering a $\langle T, p \rangle$ -policy M/G(G,G)/1 queue, and it is assumed that the system's parameters are as follows:

- the arrival rate is $\lambda = 1.0$;
- the probability that when a customer who completes FES may immediately go for SOS is $1/3$ ($\eta = 1/3$);
- startup time per busy cycle obeys 2-stage hyper-exponential with the first and second moments $E(U) = 1/2$ and $E(U^2) = 2/3$, respectively. (we choose $q_1 = 3/4$, $q_1 = 1/4$, $\gamma_{q_1} = 3$ and $\gamma_{q_2} = 1$);
- service time of FES channel is a 3-stage Erlang distribution with mean $\mu_{S_1} = 1/2$;
- service time of SOS channel is a 4-stage Erlang distribution with mean $\mu_{S_2} = 3/4$;
- FES channel and SOS channel are subject to breakdown with breakdown rate $\alpha_1 = 0.10$ and $\alpha_2 = 0.20$, respectively;

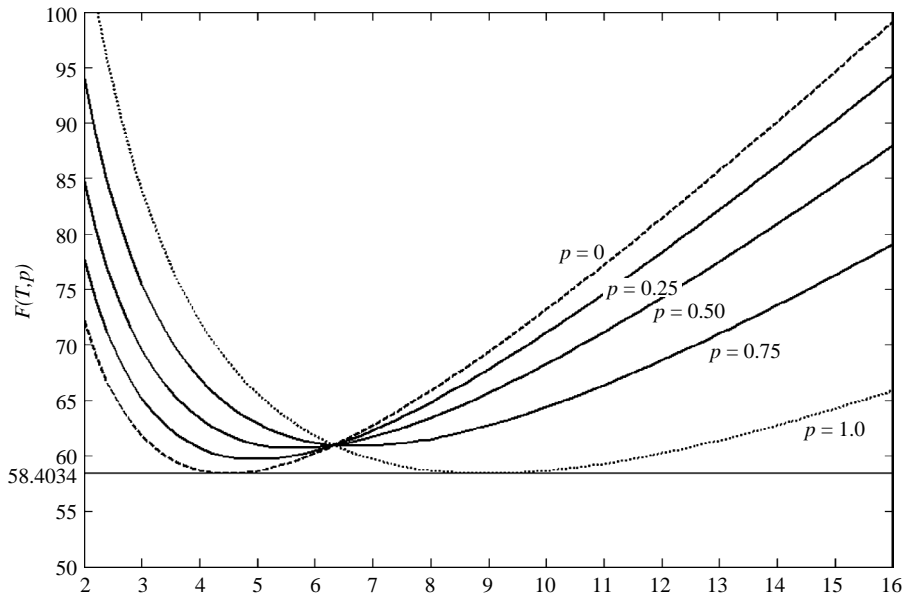


Figure 1.
The expected cost $F(T, p)$
for different values of T
under $p = 0, 0.25, 0.50,$
 0.75 and 1.0

- repair time of *FES* channel is a deterministic distribution with mean $\mu_{R_1} = 1/3$; and
- repair time of *SOS* channel is a uniform distribution over interval $[0, 2]$, say $U[0, 2]$, with mean $\mu_{R_2} = 1$.

Second optional
service and
startup

We choose $p = 0, 0.25, 0.50, 0.75, 1.0$ and fix the following cost elements $C_k = 5$, $C_f = 60$, $C_s = 100$ and $C_l = 1200$. The expected cost $F(T, p)$ for different values of T and p is shown in Figure 1. The minimum cost per unit time is \$58.403 and it is achieved at $(T^*, p^*) = (4.49, 0)$ or $(8.98, 1)$. This result is coincident with Theorem 1.

799

Concluding remarks

This paper has analyzed the optimality of the $\langle T, p \rangle$ -policy M/G/1 queue with SOS, server breakdowns and general startup times. We first derived theoretical results of some system performances for the T -policy M/G(G,G)/1 queue. Next, we utilize the convex combination property and the renewal reward theorem to obtain some important system performances for the $\langle T, p \rangle$ -policy M/G(G,G)/1 queue. The expected cost function per unit time is constructed to determine the optimal threshold values of T and p , which minimize this cost model. Analytical results for sensitivity analysis are also performed. Finally, a numerical result is presented and discussed. It is worth pointing out that the developed results still hold when η is equal to zero or one.

References

- Doshi, B.T. (1985), "A note on stochastic decomposition in a GI/G/1 queue with vacations or set-up times", *Journal of Applied Probability*, Vol. 22 No. 2, pp. 419-28.
- Doshi, B.T. (1986), "Queueing system with vacations-a survey", *Queueing Systems*, Vol. 1, pp. 29-66.
- Feinberg, E.A. and Kim, D.J. (1996), "Bicriterion optimization of an M/G/1 queue with a removable server", *Probability in the Engineering and Informational Sciences*, Vol. 10, pp. 57-73.
- Fuhrmann, S.W. and Cooper, R.B. (1985), "Stochastic decompositions in the M/G/1 queue with generalized vacations", *Operations Research*, Vol. 33, pp. 1117-29.
- Gakis, K.G., Rhee, H.K. and Sivazlian, B.D. (1995), "Distributions and first moments of the busy and idle periods in controllable M/G/1 queueing models with simple and dyadic policies", *Stochastic Analysis and Applications*, Vol. 13 No. 1, pp. 47-81.
- Gaver, D.P. (1962), "A waiting line with interrupted service, including priorities", *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 24 No. 1, pp. 73-90.
- Heyman, D.P. (1977), "The T policy for the M/G/1 queue", *Management Science*, Vol. 23, pp. 775-8.
- Ke, J.-C. (2005), "Modified T vacation policy for an M/G/1 queueing system with an unreliable server and startup", *Mathematical and Computer Modeling*, Vol. 41, pp. 1267-77.
- Ke, J.-C. (2006), "Optimal NT policies for M/G/1 system with a startup and unreliable server", *Computers & Industrial Engineering*, Vol. 50, pp. 248-62.
- Ke, J.-C. (2008a), "Two thresholds of a batch arrival queueing system under modified T vacation policy with startup and closedown", *Mathematical Methods in the Applied Sciences*, Vol. 31 No. 2, pp. 229-47.
- Ke, J.-C. (2008b), "An $M^{[k]}/G/1$ system with startup server and J additional options for service", *Applied Mathematical Modelling*, Vol. 32, pp. 443-58.

-
- Kim, D.-J. and Moon, S.-A. (2006), "Randomized control of T-policy for an M/G/1 system", *Computers & Industrial Engineering*, Vol. 51, pp. 684-92.
- Kleinrock, L. (1975), *Queueing Systems*, Vol. I, Wiley, New York, NY.
- Levy, L. and Yechiali, U. (1975), "Utilization of idle time in an M/G/1 queueing system", *Management Science*, Vol. 22, pp. 202-11.
- Madan, K.C. (2000), "An M/G/1 queue with second optional service", *Queueing Systems*, Vol. 34, pp. 37-46.
- Medhi, J. (2002), "A single server Poisson input queue with a second optional channel", *Queueing Systems*, Vol. 42, pp. 239-42.
- Sengupta, B. (1990), "A queue with service interruptions in an alternating random environment", *Operations Research*, Vol. 38 No. 2, pp. 308-18.
- Tadj, L. (2003), "A quorum queueing system under T-policy", *The Journal of the Operational Research Society*, Vol. 54, pp. 466-71.
- Takagi, H. (1991), "Queueing analysis: a foundation of performance evaluation", *Vacation and Priority Systems, Part I*, Vol. I, North-Holland, Amsterdam.
- Tang, Y.H. (1997), "A single-server M/G/1 queueing system subject to breakdowns – some reliability and queueing problem", *Microelectronics and Reliability*, Vol. 37 No. 2, pp. 315-21.
- Wang, J. (2004), "An M/G/1 queue with second optional service and server breakdowns", *Computers and Mathematics with Applications*, Vol. 47, pp. 1713-23.
- Wang, K.-H. (1997), "Optimal control of an M/E_K/1 queueing system with removable service station subject to breakdowns", *The Journal of the Operational Research Society*, Vol. 48 No. 9, pp. 936-42.
- Wang, K.-H. and Ke, J.-C. (2002), "Control policies of an M/G/1 queueing system with a removable and non-reliable server", *International Transactions in Operational Research*, Vol. 9, pp. 195-212.

Corresponding author

Kuo-Hsiung Wang can be contacted at: khwang@amath.nchu.edu.tw