Optimization of the T policy M/G/1 queue with server breakdowns and general startup times

Tsung-Yin Wang [a], Kuo-Hsiung Wang [b,*], Wen Lea Pearn [c]

[a] *Department of Accounting Information System, National Taichung Institute of Technology, Taichung 404, Taiwan*
[b] *Department of Applied Mathematics, National Chung-Hsing University, Taichung 402, Taiwan*
[c] *Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu 30050, Taiwan*

ARTICLE INFO

ABSTRACT

This paper investigates the T policy M/G/1 queue with server breakdowns, and startup times. Customers arrive at the system according to a Poisson process. Service times, repair times, and startup times are assumed to be generally distributed. The server is turned on after a fixed length of time T repeatedly until at least one customer is present in the waiting line. The server needs a startup time before starting the service. We analyze various system performance measures and develop the total expected cost function per unit time in which T is a decision variable. We determine the optimum threshold T^* and derive analytical results for sensitivity investigations. The sensitivity analysis is particularly valuable to the system analyst when evaluating future conditions. We also present extensive numerical computation for illustration purpose.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

We consider a T policy M/G/1 queue in which the server is subject to unpredictable breakdowns. It is assumed that arriving customers follow a Poisson process and the breakdown times of the server follow the negative exponential distribution. We also assume that the service times, the repair times, and the startup times obey a general distribution. After a period of time with length T, the server is turned on but is temporarily unavailable to serve the waiting customers even if there are customers in the waiting line. The server must wait until time T to begin the service. Otherwise, the server waits another period of length T and so on until at least one customer is present. When the server turns on, the server requires for the preparatory work (i.e. startup) before starting the service. Once the startup work is completed, the server immediately starts serving the waiting customers.

It is assumed that customers arrive at the system according to a Poisson process with parameter λ. The service times to the customers are independent and identically distributed (i.i.d.) random variables obeying an arbitrary distribution function $S(t)$ ($t \geq 0$) with a finite mean μ_S and a finite variance σ_S^2. The server is subject to breakdowns at any time with Poisson breakdown rate α when the server is working. When the server fails, the server is immediately sent into repair, where the repair times are i.i.d. random variables having a general distribution function $R(t)$ ($t \geq 0$) with a finite mean μ_R and a finite variance σ_R^2. Arrived customers form a single waiting line at a server based on the order of their arrivals. The server can serve only one customer at a time, and the service is independent of the arrival process. A customer who arrives and finds the server busy or broken down must wait in the queue until a server is available. Although no service occurs during the repair period of the server, customers continue to arrive following a Poisson process. If there is at least one customer in

* Corresponding author. Tel.: +886 4 22860133; fax: +886 4 22873028.
E-mail address: khwang@amath.nchu.edu.tw (K.-H. Wang).

waiting line after a period of length T, however, the server is immediately turned on (i.e. begin startup) but is temporarily unavailable to serve the waiting customers. The server needs a startup time with random length before starting the service. Again, the startup times are i.i.d. random variables obeying a general distribution function $U(t)$ ($t \geq 0$) with a finite mean μ_U and a finite variance σ_U^2. Once the startup is completed, the server begins serving the waiting customers until the system becomes empty. Service is allowed to be interrupted if the server breaks down, and the server is immediately sent into repair. Once the server is repaired, the server immediately returns to serve the customers until there is no customer in the system.

Takagi [13] studied the N policy M/G/1 queue where the server requires a startup time before starting the service. The N policy means that the server reactivates as soon as the number of customers in the system reaches N ($N \geq 1$). The server returns to service customers if there is at least one customer present in the waiting line, but when there are no customers in the system, the server waits a fixed length of time T. If after a period of length T has elapsed, there are no customers in the system, the server waits another period of length T and so on until at least one customer is present. This type of policy is called the T policy. The controllable queueing problem with T policy has been extensively investigated in the literature. A pioneering work in this field is Heyman [5] who first introduced the T policy. Without considering server breakdowns and general startup time case, the T policy M/G/1 queue was studied in [5,9,3]. Gakis et al. [3] 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. Wang and Ke [14] analyzed an M/G/1 queue with server breakdowns operating under the N policy, the T policy, and the Min (N, T) policy for which no sensitivity investigation is presented. They performed numerical comparisons of three policies which demonstrated that the optimal N policy and the optimal Min (N, T) policy are superior to T policy. We should mention that our model extended Wang and Ke's model with general startup time. Alfa and Li [1] studied the optimal (N, T) policy for an M/G/1 queue with cost structure. Hur et al. [6] investigated an M/G/1 queue under N policy, T policy, and Min (N, T) policy. They developed the steady-state distribution of the system size, the expected length of the busy and idle period, and the busy cycle for three policies, but they only determined the optimal operating N policy. Tadj [12] proposed an M/G/1 quorum queue operating under the T policy. He used probability generating function technique to obtain the steady-state system characteristics. Recently, Ke [7] examined a modified T vacation policy for an M/G/1 queue with an unreliable server and startup. Ke [7] derived the explicit formulae for various system performance measures such as the expected number of customers in the system, the expected waiting time in the queue, the expected lengths of the idle, busy, breakdown periods, and the expected length of the busy cycle, etc. We should note that in Levy and Yechiali [9], Gakis et al. [3], Tadj [12], and Ke [7], no optimal T policy with analytical results for sensitivity analysis are obtained.

Consider the manufacturing system of job-shop type, the facility can only handle one order at a time in which the fixed setup cost is incurred each time the facility is reopened. For manufacturing cost concern, the facility is only turned on when enough N orders are accumulated. This is the so-called N policy. The job-shop manufacturing system we examine here is similar to that with N policy but is controlled in another way in which the facility takes exactly one vacation when no order arrives after the facility finished all orders (T policy). In the T policy, the vacation can be utilized as a post-processing time after clearing the jobs in the system. If the manager wants to monitor continuously the queue for an arrival when the server is turned off. Although adopting the N policy is efficient in utilizing the system facilities which reduces the customer's waiting time. However, the successive monitor may result in high management cost. To reduce the management cost, the managers would prefer the T policy.

The main contribution of this paper is threefold. First, we develop various system performance measures, such as the expected number of customers in the system, the expected length of the idle, busy and breakdown periods, and the expected length of the busy cycle, etc. Second, based on the derived results, we construct the total expected cost function per unit time, including customer holding cost, the system setup cost, server on and off costs, server startup cost, and server breakdown cost. We determine the optimal threshold T^* to minimize the total expected cost. Third, we obtain analytical results, and perform a sensitivity investigation on the optimum value T^* based on changes in specific values of the system parameters (e.g., λ, μ, α, β, γ) and the cost elements (e.g., C_h, C_s, C_i, C_{sp}).

2. System performance measures

In this section, we focus mainly on developing some important system performance measures, such as (i) the expected number of customers in the system; (ii) the expected length of the idle period, the complete startup period, the busy period, and the breakdown period; (iii) the expected length of the busy cycle; and (iv) the probability that the server is idle, startup, busy and broken down.

2.1. Expected number of customers in the system

Let H be a random variable representing the completion time of a customer, which includes both the service time of a customer and the repair time of a server. Applying the well-known results of Medhi and Templeton [10], the probability generating function (p.g.f.) of the number of customers in the ordinary M/G/1 queue with server breakdowns is given by

$$G(z) = \frac{(1 - \rho_H)(1 - z)\bar{f}_H(\lambda - \lambda z)}{\bar{f}_H(\lambda - \lambda z) - z},$$

(1)

where $\rho_H = \lambda E[H]$. In addition, $E[H] = \mu_S(1 + \alpha\mu_R)$ and $E[H^2] = (1 + \alpha\mu_R)^2(\mu_S^2 + \sigma_S^2) + \alpha\mu_S(\mu_R^2 + \sigma_R^2)$ (see Wang and Ke [14]). The traffic intensity ρ_H is assumed to be less than 1. We consider a Poisson arrival process. Let ξ_i denote the elapsed time between the $(i-1)$th and the ith arriving customer. Following Ross [11], the time interval ξ_i are i.i.d. exponential random variables with mean $1/\lambda$. Let $A(t)$ denote the number of customers arriving into the system during $[0, t]$. Let A_m be the arrival time of the mth customer and let $A_m = \sum_{i=1}^{m} \xi_i$. The distribution of A_m occurring by time t is given by

$$F_{A_m}(t) = P\{A_m \leq t\} = \int_0^t \frac{\lambda(\lambda x)^{m-1}}{(m-1)!}e^{-\lambda x}dx = 1 - \sum_{k=0}^{m-1} \frac{(\lambda t)^k}{k!}e^{-\lambda t} = P\{A(t) \geq m\}.$$

It is obvious that

$$P\{A(t) = m\} = F_{A_m}(t) - F_{A_{m+1}}(t).$$

Hence, in a period of length T and when there are at least m ($m \geq 1$) customers in the system, the distribution of A_m is given

$$F_{A_m}(T) = P\{A_m \leq T\} = 1 - \sum_{k=0}^{m-1} \frac{(\lambda T)^k}{k!}e^{-\lambda T}.$$

Let $G_B(z)$ be the p.g.f. of the number of customers waiting in the queue when a busy period begins. Thus we have

$$G_B(z) = \sum_{m=0}^{\infty} z^m [F_{A_m}(T) - F_{A_{m+1}}(T)] = e^{-(1-z)\lambda T},$$

where $F_{A_0}(T) = 1$.

Let $G_u(z)$ be the p.g.f. of the number of customers arriving when given a startup period $U = u$ begins. Then we get $G_u(z) = e^{-(1-z)\lambda u}$. The Laplace–Stieltjes transform (abbreviated LST) of $G_u(z)$ is given by

$$\int_0^{\infty} G_u(z)dU(u) = \bar{f}_U[\lambda(1-z)].$$

Because the Poisson process from any point is independent of all that has previously occurred, we have $W(z) = G_B(z)\bar{f}_U[\lambda(1-z)] = e^{-(1-z)\lambda T}\bar{f}_U[\lambda(1-z)]$. For the T policy M/G/1 queue with server breakdowns and startup time, we get the complete startup period (the idle period plus the startup period) as $1 - W(z)$. Using the well-known decomposition property concerning M/G/1 vacation queue studied in [2], we obtain the p.g.f. of number of customers in the T policy M/G/1 queue with server breakdowns and general startup times as follows:

$$G_T(z) = G(z)\left[\frac{1 - W(z)}{W'(1)(1 - z)}\right], \tag{2}$$

where $G(z)$ is given in (1). Let L_T denote the expected number of customers in the T policy M/G/1 queue with server breakdowns and general startup times. Thus we obtain from (2) that

$$L_T = G_T'(z)|_{z=1} = \frac{1}{(\lambda T + \rho_U)}\left[\frac{\lambda^2 T^2}{2} + \lambda\rho_U T + \frac{\lambda^2\sigma_U^2 + \rho_U^2}{2}\right] + \rho_H + \frac{\lambda^2 E[H^2]}{2(1 - \rho_H)}, \tag{3}$$

where $\rho_U = \lambda\mu_U$.

2.2. Expected length of the idle, complete startup, busy, and breakdown periods

The idle period terminates when at least one customer arrives in system at the period T. Since the startup period starts when the idle period terminates, the complete startup period is represented by the sum of the startup period and the complete period. The server begins startup when there is at least one waiting customer at the end of the fixed period T in the system. This is called the startup period. The startup period terminates when the server starts to serve the waiting customers. Since the complete period begins when the startup period is over and terminates when the system becomes empty, the complete period is represented by the sum of the busy period and the breakdown period. The busy period is initiated when the server completes his startup and begins serving the waiting customers. During the busy period, the server may break down and sent into repair immediately. This is call the breakdown period. After the server is repaired, he returns immediately and provides service until the system is empty. Let H_O be the complete period of the ordinary M/G/1 queue with server breakdowns. Using the well-known result of Kleinrock [8, p. 213], we obtain the expected length of the complete period for the ordinary M/G/1 queue with server breakdowns as

$$E[H_O] = \frac{E[H]}{1 - \lambda E[H]} = \frac{\mu_S(1 + \alpha\mu_R)}{1 - \rho_H}. \tag{4}$$

2.2.1. Expected length of the idle period

The idle period I_T terminates when at least one customer arrives at the period T. It is obvious that

$$E[I_T] = T. \tag{5}$$

2.2.2. Expected length of the complete startup period

Let V_T represent the complete startup period for the T policy M/G/1 queue with server breakdowns and general startup times. Thus we have $V_T = H_T + U_T$, where H_T and U_T denote the complete period and the startup period, respectively. Let $\bar{f}_{V_T}(\cdot)$ be the LST of the distribution of the complete startup period of the ordinary M/G/1 queue with server breakdowns.

The following notations are used.

$F_{V_T}(\cdot)$ − distribution function of the complete startup period V_T of the T policy M/G/1 queue with server breakdowns and general startup times;

$\bar{f}_U(\cdot)$ − the LST of startup times;

$F_{H_O}(\cdot)$ − distribution function of the complete period H_O of the ordinary M/G/1 queue with server breakdowns;

$F_{H_O}^{(a+n)}(\cdot)$ − $(a + n)$ − fold convolution of $F_{H_O}(\cdot)$.

By conditioning on the length of the startup time U and the number of arrivals during U, we obtain from Gross and Harris [4, p. 277] that

$$F_{V_T/A}(x) = \int_0^x \sum_{n=0}^{\infty} P(\text{given any startup time} = t, \text{ complete startup period generated by a customer}$$

$$\text{arrival plus } n \text{ customers arrival in the complete period } H_O \text{ during } t \leq x - t)\mathrm{d}U(t)$$

$$= \int_0^x \sum_{n=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!} F_{H_O}^{(a+n)}(x - t)\mathrm{d}U(t). \tag{6}$$

Taking the LST on both sides of (6) yields

$$\bar{f}_{V_T/A}(s) = \left[\bar{f}_{H_O}(s)\right]^a \bar{f}_U \left[\lambda + s - \lambda\bar{f}_{H_O}(s)\right]. \tag{7}$$

Differentiating (7) with respect to s and then setting $s = 0$, we obtain the expected length of the complete startup period given by $A = a$ customers arriving during the period T as follows:

$$E[V_T/A = a] = (a + \lambda\mu_U) E[H] + \mu_U.$$

Using the fact that

$$E[V_T] = E[E[V_T/A]],$$

and

$$E[A] = \sum_{a=1}^{\infty} a \frac{(\lambda T)^a e^{-\lambda T}}{a!} = \lambda T,$$

we get

$$E[V_T] = E[E(V_T/A)] = E\{(A + \lambda\mu_U) E[H_O] + \mu_U\}.$$

It follows that

$$E[V_T] = E[A]E[H_O] + \lambda\mu_U E[H_O] + \mu_U = \frac{(T + \mu_U)\rho_H}{1 - \rho_H} + \mu_U. \tag{8}$$

2.2.3. Expected length of the busy and breakdown periods

Recall that $V_T = H_T + U_T$ which implies $E[V_T] = E[H_T] + E[U_T]$. Hence from (8), we obtain

$$E[H_T] = \frac{(T + \mu_U)\rho_H}{1 - \rho_H} = \frac{\mu_S(1 + \alpha\mu_R)(\lambda T + \rho_U)}{1 - \rho_H}, \tag{9}$$

and

$$E[U_T] = \mu_U. \tag{10}$$

Because the complete period is the sum of the busy period and the breakdown period, $H_T = B_T + D_T$, which implies $E[H_T] = E[B_T] + [D_T]$. From (9), we find that

$$E[B_T] = \frac{(\lambda T + \rho_U)\mu_S}{1 - \rho_H}, \tag{11}$$

and

$$E[D_T] = \frac{(\lambda T + \rho_U)\alpha\mu_S\mu_R}{1 - \rho_H}. \tag{12}$$

2.3. Expected length of the busy cycle

The busy cycle for the T policy M/G/1 queue with server breakdowns and general startup times, denoted by C_T, is the length of time from the beginning of the last idle period to the beginning of the next idle period. Since the busy cycle is the sum of the idle period (I_T), the startup period (U_T), the busy period (B_T), and the breakdown period (D_T), we get

$$E[C_T] = E[I_T] + E[U_T] + E[B_T] + E[D_T] = E[I_T] + E[V_T]. \tag{13}$$

From (5) and (8), we obtain

$$E[C_T] = \frac{T + \mu_U}{1 - \rho_H}. \tag{14}$$

2.4. Probability that the server is turned-off, startup, busy and broken down

In steady-state, let

$P_{I_T} \equiv$ probability that the server is idle;

$P_{U_T} \equiv$ probability that the server is startup;

$P_{B_T} \equiv$ probability that the server is busy;

$P_{D_T} \equiv$ probability that the server is broken down.

From (5), (10)–(12) and (14), we get

$$P_{I_T} = \frac{E[I_T]}{E[C_T]} = \frac{T(1 - \rho_H)}{T + \mu_U}, \tag{15}$$

$$P_{U_T} = \frac{E[U_T]}{E[C_T]} = \frac{\mu_U(1 - \rho_H)}{T + \mu_U}, \tag{16}$$

$$P_{B_T} = \frac{E[B_T]}{E[C_T]} = \rho, \tag{17}$$

$$P_{D_T} = \frac{E[D_T]}{E[C_T]} = \alpha \rho \mu_R, \tag{18}$$

where $\rho = \lambda \mu_S$. We demonstrate from (17) that the probability that the server is busy in the steady-state is equal to ρ.

3. The optimal T policy

We develop an expected cost function per unit time for the T policy M/G/1 queue with server breakdowns and general startup times in which T is a decision variable. Our objective is to determine the optimum value of the control parameter T, say T^*, which minimizes this cost function. Since $E[B_T]/E[C_T]$ and $E[D_T]/E[C_T]$ are not functions of the decision variable T. Therefore, we will restrict ourselves to selecting the cost elements as follows.

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

$C_s \equiv$ setup cost per busy cycle;

$C_i \equiv$ cost per unit time for keeping the server off;

$C_{sp} \equiv$ startup cost per unit time for the preparatory work of the server before starting the service.

Utilizing the definition of each cost element listed above, the expected cost function per customer per unit time is given by

$$F(T) = C_h L_T + C_s \frac{1}{E[C_T]} + C_i P_{I_T} + C_{sp} P_{U_T}. \tag{19}$$

Differentiating $F(T)$ with respect to T, we get

$$\frac{dF(T)}{dT} = \frac{C_h \left[\frac{\lambda}{2} T^2 + \lambda \mu_U T - \frac{\lambda(\sigma_U^2 - \mu_U^2)}{2} \right] - \left[C_s + (C_{sp} - C_i)\mu_U \right] (1 - \rho_H)}{(T + \mu_U)^2}. \tag{20}$$

Setting $dF(T)/dT = 0$ yields

$$T^* = -\mu_U + \sqrt{\sigma_U^2 + \frac{2 \left[C_s + (C_{sp} - C_i)\mu_U \right] (1 - \rho_H)}{C_h \lambda}}. \tag{21}$$

and since

$$\frac{\mathrm{d}^2 F(T)}{\mathrm{d}T^2}\bigg|_{T=T^*} = \frac{C_h \lambda \sigma_U^2 + 2\left[C_s + (C_{sp} - C_i)\mu_U\right](1 - \rho_H)}{(T^* + \mu_U)^3} > 0. \tag{22}$$

Thus T^* is the unique minimizer of $F(T)$.

4. Analytical results for sensitivity investigation

A system analyst may concern with varying the system parameters over a reasonable range and observing the relative change in the system performance measures. A sensitivity investigation of different system parameters (λ, μ, α, β, γ) and cost parameters (C_h, C_s, C_i, C_{sp}) levels is particularly valuable when evaluating future conditions, where $\mu_S = 1/\mu$, $\mu_R = 1/\beta$ and $\mu_U = 1/\gamma$. We can assess how robust are to system input parameters. In the following, we conduct some sensitivity investigations on the optimal value T^* based on changes in the values of system parameters and cost parameters. It implies from (21), differentiating T^* with respect to λ, we obtain

$$\frac{\partial T^*}{\partial \lambda} = \frac{-\theta_1}{\lambda^2 \sqrt{\sigma_U^2 + \frac{2\theta_1(1-\rho_H)}{\lambda}}} < 0, \tag{23}$$

where $\theta_1 = \frac{\left[C_s + (C_{sp} - C_i)\mu_U\right]}{C_h}$.

It implies from (23) that T^* decreases in λ. Differentiating T^* with respect to μ yields

$$\frac{\partial T^*}{\partial \mu} = \frac{\theta_1(1 + \alpha\mu_R)}{\mu^2 \sqrt{\sigma_U^2 + \frac{2\theta_1(1-\rho_H)}{\lambda}}} > 0. \tag{24}$$

It follows from (24) that T^* increases in μ. Similarly, differentiating T^* with respect to α and β respectively, we obtain

$$\frac{\partial T^*}{\partial \alpha} = \frac{-\theta_1 \mu_S \mu_R}{\sqrt{\sigma_U^2 + \frac{2\theta_1(1-\rho_H)}{\lambda}}} < 0, \tag{25}$$

$$\frac{\partial T^*}{\partial \mu_R} = \frac{-\theta_1 \alpha \mu_S}{\sqrt{\sigma_U^2 + \frac{2\theta_1(1-\rho_H)}{\lambda}}} < 0. \tag{26}$$

The above results imply that T^* decreases in α and μ_R, respectively. Recalling that $\mu_R = 1/\beta$, we conclude that T^* increases in β. Since σ_U^2 is a function of γ, we can see how γ affects T^* while startup time distribution is given. For special case, suppose that the startup time distribution obeys an exponential distribution with mean $\mu_U = 1/\gamma$. Substituting $\sigma_U^2 = \mu_U^2$ into (21) and then differentiating T^* with respect to μ_U, we get

$$\frac{\partial T^*}{\partial \mu_U} = -1 + \frac{\mu_U + \theta_2}{\sqrt{\mu_U^2 + 2\theta_2\mu_U + \frac{2C_s(1-\rho_H)}{C_h\lambda}}}, \tag{27}$$

where $\theta_2 = \frac{(C_{sp} - C_i)(1-\rho_H)}{C_h\lambda}$.

If $\theta_2(C_{sp} - C_i) < 2C_s$, then we have $\partial T^*/\partial \mu_U < 0$. It follows that T^* increases in γ. If $\theta_2(C_{sp} - C_i) = 2C_s$, we get $\partial T^*/\partial \mu_U = 0$. Thus T^* is independent of γ. Furthermore, if $\theta_2(C_{sp} - C_i) > 2C_s$, we have $\partial T^*/\partial \mu_U > 0$. The result implies that T^* decreases in γ. On the other hand, it can easily see from (21) that (i) T^* increases in C_s and C_{sp}; and (ii) T^* decreases in C_i and C_h.

5. Numerical comparisons

We present some numerical computations to verify the analytical results, and show how to make the decision based on minimizing the cost function $F(T)$. The sensitivity investigation concentrates mainly on the exponential startup time distribution. The cost parameters $C_s = 1000$, $C_h = 5$, $C_{sp} = 100$, $C_i = 60$ are fixed. We consider the following four cases.

Case 1: Choose $\mu = 0.5, 1.0, 1.5, 2.0$, $\alpha = 0.05$, $\beta = 3$, $\gamma = 3$, and vary the values of λ.

Case 2: Choose $\lambda = 0.2, 0.4, 0.6, 0.8$, $\alpha = 0.05$, $\beta = 3$, $\gamma = 3$, and vary the values of μ.

Case 3: Choose $\lambda = 0.5$, $\mu = 1$, $\beta = 1, 2, 3, 4$, $\gamma = 3$ and vary the values of α.

Case 4: Choose $\lambda = 0.5$, $\mu = 1$, $\alpha = 0.2, 0.4, 0.6, 0.8$, $\gamma = 3$ and vary the values of β.

Fig. 1 reveals that (i) T^* is decreasing in λ; (ii) as λ is fixed, T^* increases as μ increases; and (iii) if λ is small enough, T^* increases more quickly and the values of T^* rarely change for different values of μ. From Fig. 2, we see that (i) T^* increases in μ; (ii) if μ is small enough, T^* increases quickly; (iii) if μ is large and $\rho = \lambda/\mu$ is small enough, T^* is insensitive; and
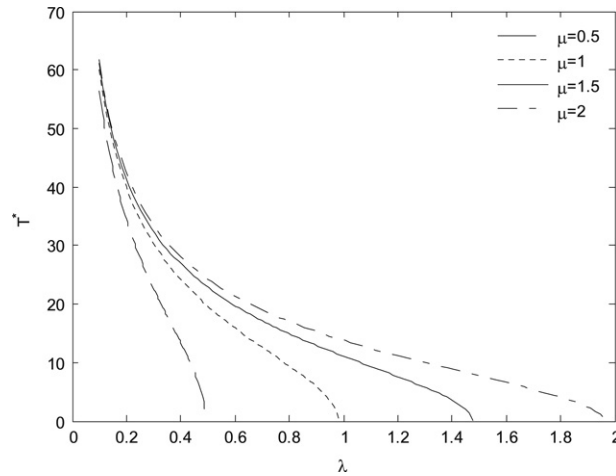
Fig. 1. Plots of (λ, T^*) with $\mu = 0.5, 1, 1.5, 2, \alpha = 0.05, \beta = 3, \gamma = 3, C_s = 1000, C_h = 5, C_{sp} = 100, C_i = 60$.
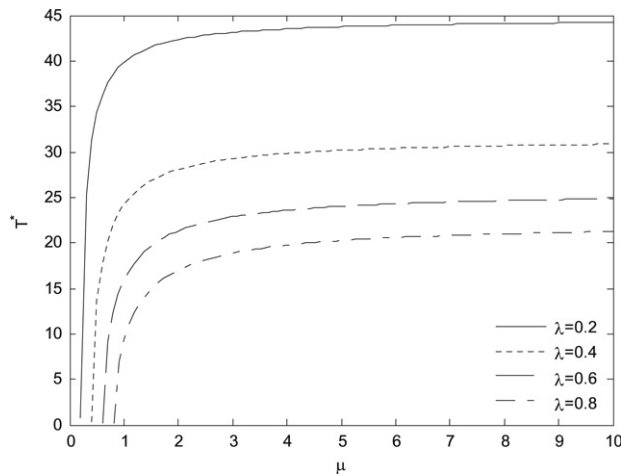

Fig. 2. Plots of (μ, T^*) with $\lambda = 0.2, 0.4, 0.6, 0.8, \alpha = 0.05, \beta = 3, \gamma = 3, C_s = 1000, C_h = 5, C_{sp} = 100, C_i = 60$.

Table 1
The optimal T^* and minimum expected cost $F(T^*)$ with various (λ, μ)

$\alpha = 0.05, \beta = 3, \gamma = 3, c_s = 1000, c_h = 5, c_{sp} = 100, c_i = 60$							
(λ, μ)	$(0.3, 0.5)$	$(0.3, 1.0)$	$(0.3, 1.5)$	$(0.3, 2.0)$	$(0.2, 1.0)$	$(0.4, 1.0)$	$(0.6, 1.0)$	$(0.8, 1.0)$
T^*	22.6241	30.3119	32.4767	33.5071	39.8499	24.1892	15.9017	9.3975
$F(T^*)$	57.8362	87.6679	97.0150	101.6107	87.9832	84.6451	72.1052	50.1233

(iv) if μ is fixed and large enough, T^* decreases as λ increases. Furthermore, the optimal value, T^*, and the corresponding minimum expected cost $F(T^*)$ are displayed in Table 1 for parameters $\alpha = 0.05, \beta = 3$, and $\gamma = 3$.

It appears from Fig. 3 that (i) T^* decreases in α; (ii) as α is fixed, the larger β has larger T^*; and (iii) T^* has an upper bound as α closes to zero. Fig. 4 reveals that (i) T^* increases in β but T^* is insensitive to β as β is large; and (ii) as β is fixed, the larger α has the smaller T^*. Furthermore, the optimal value, T^*, and the corresponding minimum expected cost $F(T^*)$ are shown in Table 2 for parameters $\lambda = 0.5, \mu = 1$, and $\gamma = 3$.

Fig. 5 indicates that (i) T^* increases in γ; and (ii) as γ is smaller than 0.4, T^* increases quickly but T^* is insensitive to γ as γ is larger than 0.4. The optimal value, T^*, and the corresponding minimum expected cost $F(T^*)$ are displayed in Table 3 for parameters $\lambda = 0.3, \mu = 1, \alpha = 0.05$, and $\gamma = 3$.

To see how T^* changes when the cost parameter changes, we set $\lambda = 0.3, \mu = 1, \alpha = 0.5, \beta = 3, \gamma = 3$, choose $C_{sp} = 100, C_i = 60$, and vary the specified values of (C_s, C_h). We observe from Table 4 that T^* increases in C_s and decreases in C_h. On the other hand, we select $C_s = 1000, C_h = 5$, and change the specified values of (C_{sp}, C_i). Table 5 reveals that T^* increases in C_{sp} and decreases in C_i, but T^* is insensitive to (C_{sp}, C_i).
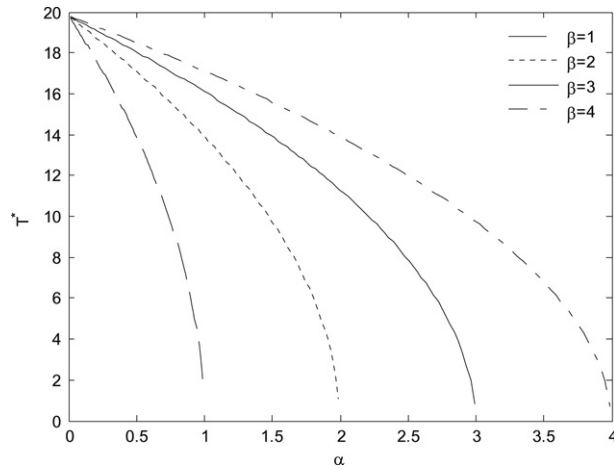
Fig. 3. Plots of (α, T^*) with $\lambda = 0.5$, $\mu = 1$, $\beta = 1, 2, 3, 4$, $\gamma = 3$, $C_s = 1000$, $C_h = 5$, $C_{sp} = 100$, $C_i = 60$.
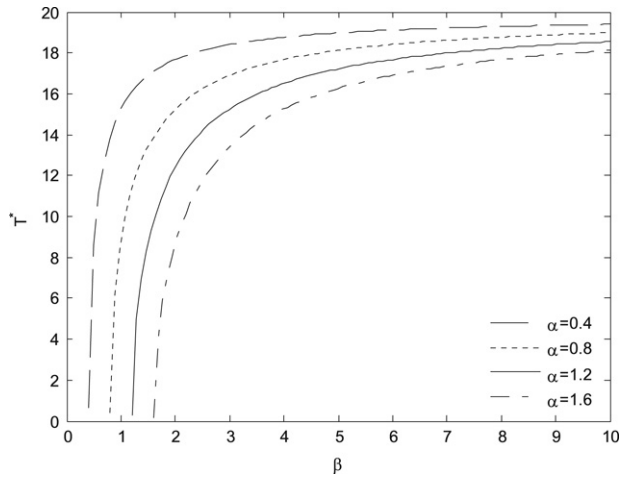

Fig. 4. Plots of (β, T^*) with $\lambda = 0.5$, $\mu = 1$, $\alpha = 0.4, 0.8, 1.2, 1.6$, $\gamma = 3$, $C_s = 1000$, $C_h = 5$, $C_{sp} = 100$, $C_i = 60$.
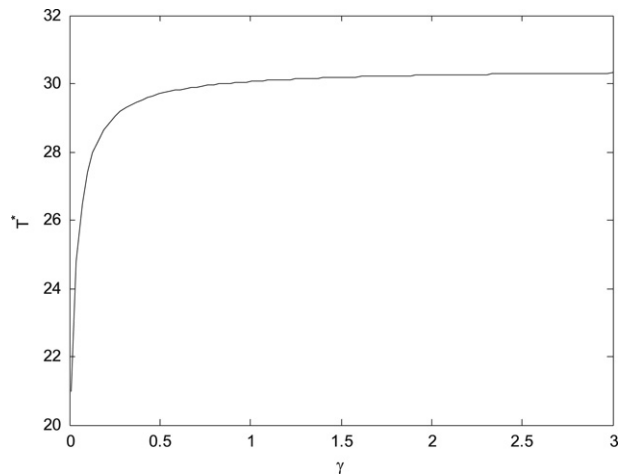

Fig. 5. Plots of (γ, T^*) with $\lambda = 0.3$, $\mu = 1$, $\alpha = 0.05$, $\beta = 3$, $C_s = 1000$, $C_h = 5$, $C_{sp} = 100$, $C_i = 60$.

Table 2
The optimal T^* and minimum expected cost $F(T^*)$ with various (α, β)

$\lambda = 0.5, \mu = 1, \gamma = 3, c_s = 1000, c_h = 5, c_{sp} = 100, c_i = 60$							
(α, β)	(0.5, 1.0)	(0.5, 2.0)	(0.5, 3.0)	(0.5, 4.0)	(0.4, 2.0)	(0.8, 2.0)	(1.2, 2.0)	(1.6, 2.0)
T^*	6.7925	8.3908	8.8621	9.0889	8.6765	7.4712	6.0420	4.1808
$F(T^*)$	59.4781	76.9783	82.4260	85.0945	80.2654	66.7251	51.7803	34.1213

Table 3
The optimal T^* and minimum expected cost $F(T^*)$ with various γ

$\lambda = 0.3, \mu = 1, \alpha = 0.05, \beta = 3, c_s = 1000, c_h = 5, c_{sp} = 100, c_i = 60$							
γ	0.4	0.8	1.2	1.6	2	2.4	2.8	3.2
T^*	7.3990	8.1910	8.4873	8.6419	8.7368	8.8009	8.8471	8.8821
$F(T^*)$	126.8279	126.2007	125.7975	125.5548	125.3955	125.2834	125.2006	125.1368

Table 4
The optimal T^* and minimum expected cost $F(T^*)$ with various (c_s, c_h)

$\lambda = 0.3, \mu = 1, \alpha = 0.05, \beta = 3, \gamma = 3, c_{sp} = 100, c_i = 60$							
(c_s, c_h)	(1000, 5)	(1000, 10)	(1000, 15)	(1000, 20)	(400, 10)	(600, 10)	(800, 10)	(900, 10)
T^*	8.8657	6.1757	4.9847	4.2753	3.8317	4.7349	5.4999	5.8470
$F(T^*)$	125.1667	159.6880	186.1419	208.4142	116.9593	133.4406	147.3823	153.7043

Table 5
The optimal T^* and minimum expected cost $F(T^*)$ with various (c_{sp}, c_i)

$\lambda = 0.3, \mu = 1, \alpha = 0.05, \beta = 3, \gamma = 3, c_s = 1000, c_h = 5$							
(c_{sp}, c_i)	(80, 20)	(80, 30)	(80, 40)	(80, 50)	(35, 25)	(45, 25)	(55, 25)	(65, 25)
T^*	8.8959	8.8808	8.8657	8.8506	8.8203	8.8355	8.8506	8.8657
$F(T^*)$	97.6411	104.4540	111.2667	118.0792	100.4285	100.5665	100.7042	100.8417

6. Conclusions

In this paper, theoretical results of some important system performance measures are derived for the T policy M/G/1 queue with server breakdowns and general startup times. We established a cost function to determine the optimal threshold T numerically to minimize the expected cost. We performed a sensitivity investigation to discuss how the optimal threshold T can be affected by the changes of the input parameters (or cost parameters) in the studied queueing service model. The sensitivity investigation is particularly valuable to the system analyst when evaluating future condition of the system.

References

[1] A.S. Alfa, W. Li, Optimal (N, T)-policy for M/G/1 system with cost structure, Perform. Eval. 42 (2000) 265–277.
[2] S.W. Fuhrmann, R.B. Cooper, Stochastic decompositions in the M/G/1 queue with generalized vacations, Oper. Res. 33 (1985) 1117–1129.
[3] K.G. Gakis, H.K. Rhee, B.D. Sivazlian, Distributions and first moments of the busy and idle periods in controllable M/G/1 queueing models with simple and dyadic policies, Stoch. Anal. Appl. 13 (1995) 47–81.
[4] D. Gross, C.M. Harris, Fundamentals of Queueing Theory, Wiley, New York, 1985.
[5] D.P. Heyman, The T policy for the M/G/1 queue, Mgmt. Sci. 23 (1977) 775–778.
[6] S. Hur, J. Kim, C. Kang, An analysis of the M/G/1 system with N and T policy, Appl. Math. Modelling 27 (2003) 665–675.
[7] J.-C. Ke, Modified T vacation policy for M/G/1 queueing system with an unreliable server and startup, Math. Comput. Modelling 41 (2005) 1267–1277.
[8] L. Kleinrock, Queueing Systems: Theory, vol. I, John Wiley & Sons, Inc., New York, 1975.
[9] Y. Levy, U. Yechiali, Utilization of idle time in an M/G/1 queueing system, Mgmt. Sci. 22 (1975) 202–211.
[10] J. Medhi, J.G.C. Templeton, A Poisson input queue under N-policy and with a general start up time, Comput. Oper. Res. 19 (1992) 35–41.
[11] S.M. Ross, Introduction to Probability Models, Academic Press, 2003.
[12] L. Tadj, On an M/G/1 quorum queueing system under T policy, J. Opl. Res. Soc. 54 (2003) 466–471.
[13] H. Takagi, M/G/1/K queues with N-policy and setup times, Queueing Syst. 14 (1993) 79–98.
[14] K.-H. Wang, J.-C. Ke, Control policies of an M/G/1 queueing system with a removable and non-reliable server, Intl. Trans. in Opl. Res. 9 (2002) 195–212.