academicJournals

Vol. 7(13), pp. 1035-1048, 7 April, 2013
DOI: 10.5897/AJBM11.814
ISSN 1993-8233 © 2013 Academic Journals
http://www.academicjournals.org/AJBM

African Journal of Business Management

Full Length Research Paper

Cost analysis of a discrete-time queue with a single vacation and randomized activation

Tsung-Yin Wang[1]*, Jau-Chuan Ke[2], Chuen-Horng Lin[3] and Fu-Min Chang[4]

[1]Department of Accounting Information, National Taichung Institute of Technology, Taichung 404, Taiwan, ROC.
[2]Department of Applied Statistics, National Taichung Institute of Technology, Taichung 404, Taiwan, ROC.
[3]Graduate School of Computer Science and Information Technology, National Taichung Institute of Technology, Taichung 404, Taiwan, ROC.
[4]Department of Finance, Chaoyang University of Technology, Taichung 404, Taiwan, ROC.

Accepted 28 February, 2012

This paper considers a discrete-time Geo/G/1 queue, in which the server operates a single vacation at end of each consecutive service period. After all the messages are served in the queue exhaustively, the server immediately leaves for a vacation. Upon returning from the vacation, the server inspects the queue length. If there are some messages waiting in the queue, the server either resumes serving the waiting messages (with probability *p*) or remains idle in the system (with probability 1-*p*) until the next message arrives; and if no message presents in the queue, the server stays dormancy in the system until at least one message arrives. Using the generating functions technique, the system state evolution is analyzed. The probability generating functions of the system size distributions in various states are obtained. Some system characteristics of interest are also derived. With the vacation of fixed length time (say *T*), the long run average cost function per unit time is analytically developed to determine the joint optimal values of *T* and *p* at a minimum cost.

Key words: Cost, busy period, discrete time queue, markov chain, randomized vacation.

INTRODUCTION

Starting from Levy and Yechiali (1975), the modelling analysis for the queueing systems with vacations has been done by a considerable amount of work in the past. A comprehensive and excellent study on the vacation models, including some applications such as production/ inventory system and communication/ computer systems, can be found in Doshi (1986), Takagi (1991), and Tian and Zhang (2006).

On the other hand, along with the advent of computer and communication technologies, the analysis of discrete-time queueing systems has received more attention in the scientific literatures over the past years (Meisling, 1958; Hunter, 1983; Bruneel and Kim, 1993; Takagi, 1993; Woodward, 1994). The reason for this is that discrete-time systems are more appropriate than their continuous-time counterparts in their applicability for the study of many computer and communication systems applications in which time is divided into fixed-length time intervals ('slots'). The applications to communication and computer systems include asynchronous transfer mode multiplexers in the broadband integrated services digital network, slotted carrier-sense multiple access protocols, and time-division multiple access schemes. An excellent and complete study on discrete-time queueing systems with vacations has been presented by Takagi (1993). Zhang and Tian (2001) investigated a Geo/G/1 queue with multiple adaptive vacations. Tian and Zhang (2002) analyzed a GI/Geo/1 queueing system with multiple vacations by matrix-geometric solution method and Li and Tian (2007) used the same method to study a GI/Geo/1 queueing system with working vacation and vacation interruption.

*Corresponding author. E-mail: ciwang@ntit.edu.tw.

Zhang and Tian (2001), and Li and Tian (2002), they gave the stochastic results for the queue length and waiting time. We should note that in Zhang and Tian (2001), and Li and Tian (2002), no optimal vacation policies for sensitivity analysis are obtained.

In this paper, a discrete-time Geo/G/1 system with a single vacation policy with randomized activation (namely <*V*, *p*> policy) was considered. The <*V*, *p*> policy is performed under the following conditions: (i) the server leaves for a single vacation when the system is empty, (ii) if the server returns from the vacation and at least one message is waiting in the queue, the server may either activates with probability *p* or stays dormancy in the system with probability $q(=1-p)$,and (iii) when the server is dormant in the system, he activates to serve the waiting messages as the next message arrives.

To the best of our knowledge, the discrete-time Geo/G/1 queue with <*V*, *p*> policy has not been studied. Such a model has a potential application in wireless local area networks (WLANs). Access Points (APs) are specially configured nodes on WLANs and act as a central transmitter and receiver of WLAN radio signals. To keep the APs functioning well, some maintenance activities are needed. For example, virus scan is an important maintenance activity for the APs. It can be performed when the AP is idle and be programmed to perform on a regular basis. After finishing the maintenance activity, AP can enter the sleep mode when there is no radio signal to be transmitted for power saving. It can also enter the sleep mode after finishing the some kinds of maintenance activities such as refreshing AP current status. AP will awake from sleep mode and begin to serve when the new radio signal arrives.

In this study, we first investigate the Markov chain, the probability generating functions and probability distributions of queue length in various server states, including their main expectation. Next, we derive the turned-off period, turned-on (busy) period of the server and the waiting time distribution in the queue. Finally, we develop a cost model with a fixed positive integer *V* (say *T*), and the joint optimal threshold values (T^*, p^*) are determined which minimize the cost.

MODEL DESCRIPTION

The study considers a discrete-time queue with a single-server where the time is divided into constant length intervals (called slots). It is well-known that the probability of an arrival and a departure occurring simultaneously is not zero in discrete time. This probability is positive in the discrete-time setting. That is why the order of the arrivals and departures must be stated. There are two different laws: if the arrivals precede the departures (late arrival system (LAS)) and if the departures precede the arrivals (early arrival system (EAS)). These concepts and other related ones can be found in Takagi (1993). We adopt the LAS policy in the present model (Figure 1).

Messages arrive according to a Bernoulli process with rate λ, that is, λ (respectively, $\overline{\lambda}=1-\lambda$) is the probability that a message arrives (respectively, does not arrive) in each slot. The service times of the messages are independent and identically distributed according to a general probability mass function $\{b_i\}_{i=1}^{\infty}$ with probability generating function (pgf) $B(u)=\sum_{i=1}^{\infty} b_i u^i$ and *j*th factorial moments B_j. After all the messages are served in the queue exhaustively, the server operates a <*V*, *p*> policy. As soon as the system becomes empty, the server immediately takes a single vacation, where the vacation time is a discrete random variable, denoted by *V*, with probability mass function $\{v_i\}_{i=1}^{\infty}$ having pgf $V(u)=\sum_{i=1}^{\infty} v_i u^i$ and *j*th factorial moments V_j. At the vacation completion instant, the server checks the system to see if there is any waiting message and decides the action to take one of the following two cases according to the state of the system:

Case 1: If there is any message waiting in the queue, the server will resume serving the queue with probability *p* or to stay dormancy in the system with probability $q\ (=1-p)$ until at least one message arrives.

Case 2: If there is no message waiting in the queue, the server remains idle in the system until the next message arriving.

Arriving messages form a single waiting line based on the order of their arrivals; that is, they are queued according to the first-come, first-served (FCFS) discipline. The server can serve only one message at a time. If the server is busy, arriving message has to wait in the queue until the server is available. All messages arriving to the system are assumed to be eventually served, that is, $\lambda B_1 < 1$. Furthermore, various stochastic processes involved in the system are independent of each other.

THE ANALYSIS: MARKOV CHAIN AND PROBABILITY GENERATING FUNCTION

At time n^+, the state of the system is described by the process $(\Phi^{(n)}, L^{(n)}, \xi^{(n)})$. $\Phi^{(n)}$ denotes the state of the server, where it can be 0, 1, or 2 representing the server on vacation, idle, or busy, respectively. As Takagi (1993), $L^{(n)}$ is the number of messages in the system. If $\Phi^{(n)}=0$, then $\xi^{(n)}$ denotes the remaining vacation time. If $\Phi^{(n)}=2$, then $\xi^{(n)}$ represents the remaining service time. The sequence of triplets

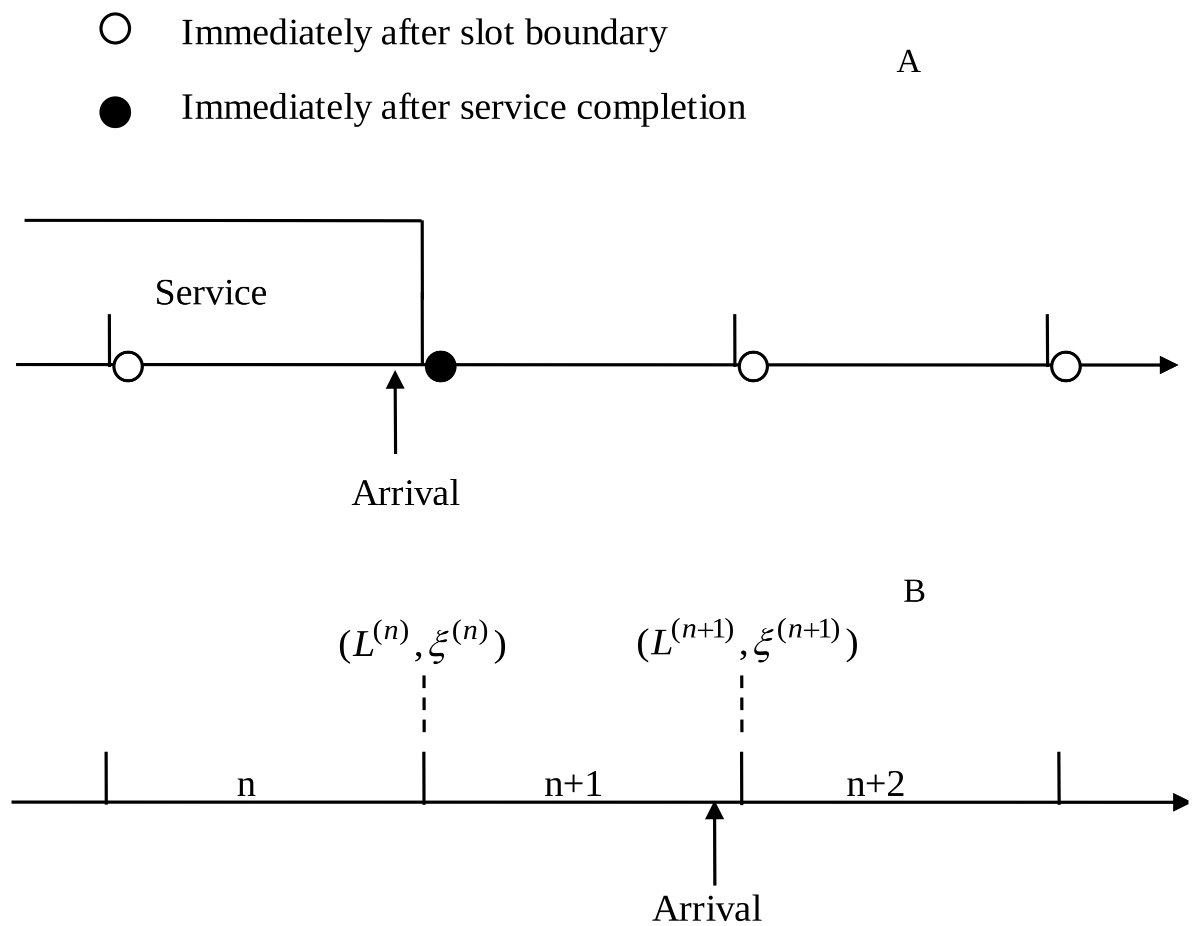

Figure 1. A: Late arrival model. B: Queue size and remaining service time observed immediately after slot boundaries.

$\{(\Phi^{(n)}, L^{(n)}, \xi^{(n)}); n = 0,\ 1,\ 2, \cdots\}$ is a Markov chain whose state space is as $\{(j,k,i) : j = 0,2, k \geq 0, i \geq 1; (1,k) : k \geq 0\}$.

Let us define the following limiting probabilities:

$\tilde{\pi}_{k,i} = \lim_{n\to\infty} \Pr[\Phi^{(n)} = 0, L^{(n)} = k,\ \xi^{(n)} = i]$,

$k \geq 0,\ i \geq 1,\qquad \Omega_k = \lim_{n\to\infty} \Pr[\Phi^{(n)} = 1,\ L^{(n)} = k],$

$k \geq 0$, $\pi_{k,i} = \lim_{n\to\infty} \Pr[\Phi^{(n)} = 2, L^{(n)} = k, \xi^{(n)} = i],\quad k \geq 1, i \geq 1$.

The Kolmogorov equations for the stationary distribution are given by

$$\tilde{\pi}_{0,i} = \pi_{1,1}\bar{\lambda} v_i + \tilde{\pi}_{0,i+1}\bar{\lambda},\ i \geq 1 \tag{1}$$

$$\tilde{\pi}_{k,i} = \tilde{\pi}_{k,i+1}\bar{\lambda} + \tilde{\pi}_{k-1,i+1}\lambda,\ k \geq 1,\ i \geq 1 \tag{2}$$

$$\Omega_0 = \tilde{\pi}_{0,1}\bar{\lambda} + \Omega_0\bar{\lambda} \tag{3}$$

$$\Omega_k = q\tilde{\pi}_{k,1}\bar{\lambda} + q\tilde{\pi}_{k-1,1}\lambda + \Omega_k\bar{\lambda},\ k \geq 1 \tag{4}$$

$$\pi_{1,i} = \Omega_0\lambda b_i + p\tilde{\pi}_{0,1}\lambda b_i + p\tilde{\pi}_{1,1}\bar{\lambda} b_i + \pi_{1,i+1}\bar{\lambda} + \pi_{2,1}\bar{\lambda} b_i + \pi_{1,1}\lambda b_i,\ i \geq 1 \tag{5}$$

$$\pi_{k,i} = \Omega_{k-1}\lambda b_i + p\tilde{\pi}_{k-1,1}\lambda b_i + p\tilde{\pi}_{k,1}\bar{\lambda} b_i + \pi_{k,i+1}\bar{\lambda} + \pi_{k+1,1}\bar{\lambda} b_i + \pi_{k-1,i+1}\lambda + \pi_{k,1}\lambda b_i,\ k \geq 2,\ i \geq 1. \tag{6}$$

To resolve (1) to (8), we use the following generating functions:

$$G_V(u,z) = \sum_{k=0}^{\infty}\sum_{i=1}^{\infty} z^k u^i \tilde{\pi}_{k,i},\ G_I(z) = \sum_{k=0}^{\infty} z^k \Omega_k,\ V(u) = \sum_{i=1}^{\infty} u^i v_i, \quad (1)$$

$$B(u) = \sum_{i=1}^{\infty} u^i b_i,\ \ G_B(u,z) = \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} z^k u^i \pi_{k,i},\ \varphi_V(z) = \sum_{k=0}^{\infty} z^k \tilde{\pi}_{k,1},$$

$$\varphi_B(z)=\sum_{k=1}^{\infty} z^k \pi_{k,1} \ (|z|\le 1 \text{ and } |u|\le 1),$$

And the normalization condition

$$G_V(1,1)+G_I(1)+G_B(1,1)=1.$$

Multiplying (1) and (2) by z^k and summing over k after multiplying (1) and (2) by u^i and summing over i, it finally yields:

$$\frac{u-(\bar{\lambda}+\lambda z)}{u}G_V(u,z)=\bar{\lambda}\pi_{1,1}V(u)-(\bar{\lambda}+\lambda z)\varphi_V(z)\cdot \quad (7)$$

Inserting $u=(\bar{\lambda}+\lambda z)$ in (7), we obtain

$$\varphi_V(z)=\frac{\bar{\lambda}V(\bar{\lambda}+\lambda z)}{(\bar{\lambda}+\lambda z)}\pi_{1,1}. \quad (8)$$

Substituting (8) into (7), we get

$$G_V(u,z)=\frac{\bar{\lambda}u\left[V(u)-V(\bar{\lambda}+\lambda z)\right]}{u-(\bar{\lambda}+\lambda z)}\pi_{1,1}. \quad (9)$$

Multiplying (3) and (4) by z^k and summing over k, we obtain

$$\lambda G_I(z)=p\bar{\lambda}\tilde{\pi}_{0,1}+q(\bar{\lambda}+\lambda z)\varphi_V(z). \quad (10)$$

Substituting (8) into (10), it gives

$$G_I(z)=\frac{\bar{\lambda}}{\lambda}\left(p\tilde{\pi}_{0,1}+qV(\bar{\lambda}+\lambda z)\pi_{1,1}\right). \quad (11)$$

Multiplying (5) and (6) by z^k and summing over k after multiplying (5) and (6) by u^i and summing over i, we have

$$\frac{u-(\bar{\lambda}+\lambda z)}{u}G_B(u,z)=\lambda zB(u)G_I(z)+p(\bar{\lambda}+\lambda z)B(u)\varphi_V(z)-p\bar{\lambda}B(u)\tilde{\pi}_{0,1}$$

$$+\left[\frac{B(u)-z}{z}\right](\bar{\lambda}+\lambda z)\varphi_B(z)-\bar{\lambda}B(u)\pi_{1,1}. \quad (12)$$

Letting $u=(\bar{\lambda}+\lambda z)$ in (12), it gives

$$\varphi_B(z)=\frac{\bar{\lambda}zB(\bar{\lambda}+\lambda z)\left\{p(1-z)\tilde{\pi}_{0,1}+\left[1-(p+qz)V(\bar{\lambda}+\lambda z)\right]\pi_{1,1}\right\}}{(\bar{\lambda}+\lambda z)\left[B(\bar{\lambda}+\lambda z)-z\right]} \quad (13)$$

Substituting (8), (11), (13) into (12), we obtain

$$G_B(u,z)=\frac{\bar{\lambda}uz\left[B(u)-B(\bar{\lambda}+\lambda z)\right]\left\{p(1-z)\tilde{\pi}_{0,1}+\left[1-(p+qz)V(\bar{\lambda}+\lambda z)\right]\pi_{1,1}\right\}}{\left[u-(\bar{\lambda}+\lambda z)\right]\cdot\left[B(\bar{\lambda}+\lambda z)-z\right]} \quad (14)$$

Letting *u* = 1 in (9) and (14) respectively, it yields

$$G_V(1,z)=\frac{\bar{\lambda}\left[1-V(\bar{\lambda}+\lambda z)\right]}{1-(\bar{\lambda}+\lambda z)}\pi_{1,1}, \quad (15)$$

and

$$G_B(1,z)=\frac{\bar{\lambda}z\left[1-B(\bar{\lambda}+\lambda z)\right]\left\{p(1-z)\tilde{\pi}_{0,1}+\left[1-(p+qz)V(\bar{\lambda}+\lambda z)\right]\pi_{1,1}\right\}}{\left[1-(\bar{\lambda}+\lambda z)\right]\cdot\left[B(\bar{\lambda}+\lambda z)-z\right]} \quad (16)$$

Let S(z) be the probability generating function of the number of messages in the system. Since $S(z)=G_V(1,z)+G_I(z)+G_B(1,z)$, it follows from (11), (15) and (16) that

$$S(z)=\frac{\bar{\lambda}B(\bar{\lambda}+\lambda z)\left\{p(1-z)\tilde{\pi}_{0,1}+\left[1-(p+qz)V(\bar{\lambda}+\lambda z)\right]\pi_{1,1}\right\}}{\lambda\left[B(\bar{\lambda}+\lambda z)-z\right]} \quad (17)$$

The derivation of $\tilde{\pi}_{0,1}$ and $\pi_{1,1}$

Differentiating $G_V(u,z)=\sum_{k=0}^{\infty}\sum_{i=1}^{\infty} z^k u^i \tilde{\pi}_{k,i}$ with respect to *u* and then setting $u=z=0$, we obtain

$$\left.\frac{\partial}{\partial u}G_V(u,z)\right|_{u=z=0}=\left.\sum_{k=0}^{\infty}\sum_{i=1}^{\infty} iz^k u^{i-1}\tilde{\pi}_{k,i}\right|_{u=z=0}=\tilde{\pi}_{0,1}. \quad (18)$$

By proceeding similar manner with Equation (9) yields

$$\left.\frac{\partial}{\partial u}G_V(u,z)\right|_{z=u=0}=V(\bar{\lambda})\pi_{1,1}. \quad (19)$$

It follows from (18) and (19) that

$$\pi_{1,1}=\frac{\tilde{\pi}_{0,1}}{V(\bar{\lambda})}. \quad (20)$$

Using the normalization condition $S(1)=1$ and Equation (20), it yields

$$\tilde{\pi}_{0,1} = \frac{\lambda(1-\rho)V(\bar{\lambda})}{\bar{\lambda}\left[q+\lambda V_1 + pV(\bar{\lambda})\right]}, \tag{21}$$

and

$$\pi_{1,1} = \frac{\tilde{\pi}_{0,1}}{V(\bar{\lambda})} = \frac{\lambda(1-\rho)}{\bar{\lambda}\left[q+\lambda V_1 + pV(\bar{\lambda})\right]}. \tag{22}$$

Substitution $\tilde{\pi}_{0,1}$ and $\pi_{1,1}$ to Equation (17) gives

$$S(z) = S_0(z)\cdot\frac{\left\{p(1-z)V(\bar{\lambda}) + \left[1-(p+qz)V(\bar{\lambda}+\lambda z)\right]\right\}}{(1-z)\left[q+\lambda V_1 + pV(\bar{\lambda})\right]}, \tag{23}$$

Where $S_0(z)$ is the pgf for the classical Geo/G/1 queue without vacation and

$$S_0(z) = \frac{(1-z)(1-\rho)B(\bar{\lambda}+\lambda z)}{\left[B(\bar{\lambda}+\lambda z)-z\right]}.$$

Remark: As p = 1, $S(z)$ can be simplified into

$$S(z) = S_0(z)\cdot\frac{\left\{(1-z)V(\bar{\lambda}) + \left[1-V(\bar{\lambda}+\lambda z)\right]\right\}}{(1-z)\left[V(\bar{\lambda})+\lambda V_1\right]},$$

This is referred to the Geo/G/1 queue with a single vacation.

Stationary distribution of the server state

Let us define the following probabilities

$P_V \equiv$ The probability that the server is on vacation;

$P_I \equiv$ The probability that the server is idle;

$P_{on} \equiv$ The probability that the server is turned on (working);

$P_{off} \equiv$ The probability that the server is turned off (vacation or idle).

Substituting (21) and (22) into (11), (15) and (16) and then setting z = 1, it yields

$$P_V = \lim_{z\to 1} G_V(1,z) = \frac{\lambda(1-\rho)V_1}{\left[q+\lambda V_1 + pV(\bar{\lambda})\right]}. \tag{24}$$

By proceeding similar manner with $G_I(z)$ and $G_B(1,z)$, it finally yields

$$P_I = G_I(1) = \frac{(1-\rho)\left[q+pV(\bar{\lambda})\right]}{\left[q+\lambda V_1 + pV(\bar{\lambda})\right]}, \tag{25}$$

and

$$P_{on} = \rho. \tag{26}$$

From (24) and (25), the study obtains

$$P_{off} = P_V + P_I = 1-\rho. \tag{27}$$

The probability of system empty

Now, we want to find that the probability of system empty. Inserting u = 1 and z = 0 in (9) and using (22), we obtain:

$$\sum_{i=1}^{\infty}\tilde{\pi}_{0,i} = \frac{(1-\rho)\left[1-V(\bar{\lambda})\right]}{\left[q+\lambda V_1 + pV(\bar{\lambda})\right]}. \tag{28}$$

It follows from (3) and (21) that we have

$$\Omega_0 = \frac{(1-\rho)V(\bar{\lambda})}{\left[q+\lambda V_1 + pV(\bar{\lambda})\right]}. \tag{29}$$

Thus the probability of system empty is as

$$P_0 = \sum_{i=1}^{\infty}\tilde{\pi}_{0,i} + \Omega_0 = \frac{(1-\rho)}{\left[q+\lambda V_1 + pV(\bar{\lambda})\right]}. \tag{30}$$

The expected number of messages in the system

Differentiating $S(z)$ in (23), we note that the numerator and denominator are both 0. The study applies L'Hospital's rule twice and finally find the expected number of messages in the system given by,

$$L_{V,p} = L + \frac{2q\lambda V_1 + \lambda^2 V_2}{2\left[q+\lambda V_1 + pV(\bar{\lambda})\right]}, \tag{31}$$

Where $L = \rho + \dfrac{\lambda^2 B_2}{2(1-\rho)}$ is the expected number of messages in the system for the classical Geo/G/1 queue without vacations (Takagi, 1993: 6).

THE TURNED-OFF PERIOD AND TURNED-ON PERIOD

This section studies the turned-off period (which is comprised of vacation period and idle period) and turned-on period (busy period). Let us define the following terminology: A vacation period starts at the departure instant of a message which leaves the system empty and terminates at returning instant from a vacation.

An idle period starts the end of a vacation and terminates at the end of the next succeeding slot during which an arrival occurs. A busy period starts at the beginning of a service and terminates when a service is completed and the system is empty.

The turned-off period

According to the definition, we have

1. The joint pgf for the length of a vacation and the probability that no message arrives during that vacation is given by $\sum_{i=1}^{\infty} u^i v_i \bar{\lambda}^i = V(\bar{\lambda} u)$.
2. The joint pgf for the length of a vacation and the probability that at least one message arrives during that vacation is given by $V(u) - V(\bar{\lambda} u)$.
3. The pgf for the length of an idle period is given by $I(u) = \lambda u / (1 - \bar{\lambda} u)$.

From the results listed previously, the pgf of the server turned-off period is given by

$$I_V(u) = p\left[V(u) - V(\bar{\lambda} u)\right] + V(\bar{\lambda} u) I(u) + q\left[V(u) - V(\bar{\lambda} u)\right] I(u), \quad (32)$$

Which leads to the expected length of the turned-off period as

$$E[S_{off}] = I_V'(1) = V_1 + \frac{pV(\bar{\lambda}) + q}{\lambda}. \quad (33)$$

The turned-on (busy) period

Let Ψ be the pgf of busy period of classical Geo/G/1 with late arrive delay access. From Takagi (1993), we have

$$\Psi(z) = B(\lambda z \Psi(z) + \bar{\lambda} z) \quad (34)$$

The busy period begins as one of the following three cases:

Case 1: *j* messages arrive during the vacation period which vacation time is *k* slots. After the vacation completion instant, the server begins service with probability *p*. Such event occurs with probability $pv_k^{(j)} = pC_j^k \lambda^j \bar{\lambda}^{k-j} v_k$, ($k = 1, 2 \cdots$, $j = 1, 2, \cdots k$).

Case 2: *j* messages arrive during the vacation period which vacation time is *k* slots. At the end of the vacation, the server remains idle in the system with probability *q*. In this case, the server begins providing service as next message arrives. Such event occurs with probability

$$qv_k^{(j)} \sum_{m=1}^{\infty} \bar{\lambda}^{m-1} \lambda = qC_j^k \lambda^j \bar{\lambda}^{k-j} v_k,$$

($k = 1, 2 \cdots$, $j = 1, 2, \cdots k$).

Case 3: No message arrives during the vacation period which vacation time is *k* slots. In this case, the server starts providing service as a message arrives. Such event occurs,

$$\bar{\lambda}^k v_k \sum_{m=1}^{\infty} \bar{\lambda}^{m-1} \lambda = \bar{\lambda}^k v_k, k = 1, 2 \cdots.$$

According to the definition, the pgf of the sub-busy period is extended by case 1 as

$$\sum_{k=1}^{\infty} \sum_{j=1}^{k} pv_k^{(j)} \sum_{n=1}^{\infty} b_n z^n \sum_{l=0}^{n} C_l^n \lambda^l \bar{\lambda}^{n-l} \left(\Psi(z)\right)^{l+j-1}.$$

The pgf of the sub-busy period extended by case 2 is given by

$$\sum_{k=1}^{\infty} \sum_{j=1}^{k} qv_k^{(j)} \sum_{n=1}^{\infty} b_n z^n \sum_{l=0}^{n} C_l^n \lambda^l \bar{\lambda}^{n-l} \left(\Psi(z)\right)^{l+j}.$$

The pgf of the sub-busy period extended by case 3 is given by

$$\sum_{k=1}^{\infty} \bar{\lambda}^k v_k \sum_{n=1}^{\infty} b_n z^n \sum_{l=0}^{n} C_l^n \lambda^l \bar{\lambda}^{n-l} \left(\Psi(z)\right)^{l}.$$

From the results listed previously, the pgf of the busy period for the <*V*, *p*> policy Geo/G/1 queueing system is given by

$$\bar{\Psi}(z) = \sum_{k=1}^{\infty} \sum_{j=1}^{k} pv_k^{(j)} \sum_{n=1}^{\infty} b_n z^n \sum_{l=0}^{n} C_l^n \lambda^l \bar{\lambda}^{n-l} \left(\Psi(z)\right)^{l+j-1} + \sum_{k=1}^{\infty} \sum_{j=1}^{k} qv_k^{(j)} \sum_{n=1}^{\infty} b_n z^n \sum_{l=0}^{n} C_l^n \lambda^l \bar{\lambda}^{n-l} \left(\Psi(z)\right)^{l+j}$$

$$+ \sum_{k=1}^{\infty} \bar{\lambda}^k v_k \sum_{n=1}^{\infty} b_n z^n \sum_{l=0}^{n} C_l^n \lambda^l \bar{\lambda}^{n-l} \left(\Psi(z)\right)^{l}$$

$$= B\left(\bar{\lambda} z + \lambda z \Psi(z)\right) \left\{ \left(\frac{p}{\Psi(z)} + q \right) \left[V(\bar{\lambda} + \lambda \Psi(z)) - V(\bar{\lambda}) \right] + V(\bar{\lambda}) \right\}, \quad (35)$$

Which implies the expected length of the turned-on period (busy period) is given by

$$E[S_{on}] = \overline{\Psi}'(1) = \frac{\rho\left[q + \lambda V_1 + pV(\overline{\lambda})\right]}{\lambda(1-\rho)}. \qquad (36)$$

From (33) and (36), we obtain the expected length of busy cycle

$$E[C_V] = E[S_{off}] + E[S_{on}] = \frac{\left[q + \lambda V_1 + pV(\overline{\lambda})\right]}{\lambda(1-\rho)}. \qquad (37)$$

WAITING TIME IN THE QUEUE

The waiting time in the queue (measured in slots) in the slot n +1 as the time that a message would wait in the corresponding system if it arrived in the slot n +1 (Takagi, 1993). Thus, the waiting time in the system (measured in slots) in the slot n +1 is the sum of the waiting time in the queue in the slot n +1 plus the service time. Let us define the following pgfs:

$W_V(z) \equiv$ The pgf of the waiting time in the queue of a test message (that arrives in the slot n +1) when server is on vacation.
$W_I(z) \equiv$ The pgf of the waiting time in the queue of a test message (that arrives in the slot n +1) when server is in idle period.
$W_B(z) \equiv$ The pgf of the waiting time in the queue of a test message (that arrives in the slot n +1) when server is busy.

$W_Q(z) \equiv$ The pgf of waiting time in the queue of a test message that arrives in the slot n +1.

For the Geo/G/1 system with <V, p> policy, an arrival may occurs as one of the following three cases:

Case 1: The test message that arrives while the server is on vacation and find k message($k \geq 0$) in the system: (i) while the vacation is just end, the server is switched to busy period with probability p, the test message must wait the service time of the preceding k messages; and (ii) while the vacation is just end, the server is switched to idle period with probability q until the next message arrives, the test message must wait the next message arriving plus the service time of the preceding k messages.
Case 2: There are exactly k messages in the queue and the server is idle in the system when the message arrives.
Case 3: The test message that arrives while the server is busy and finds k messages in the system. In this case, the waiting time in the queue of the message consists of:

(i) The remaining service time of the message being served at time n^+; and
(ii) The service time of the k-1 messages in the queue at time n^+.

From Case 1 yields,

$$W_V(z) = \frac{1}{(1-\rho-P_I)}\left[p\sum_{k=0}^{\infty}\sum_{i=1}^{\infty}\tilde{\pi}_{k,i}z^{i-1}\left[B(z)\right]^k + q\sum_{k=0}^{\infty}\sum_{i=1}^{\infty}\tilde{\pi}_{k,i}z^{i-1}[B(z)]^k\left(\frac{\lambda z}{1-\overline{\lambda}z}\right)\right]$$
$$= \frac{1}{(1-\rho-P_I)}\left[\frac{p}{z}G_V(z,B(z)) + \frac{q}{z}G_V(z,B(z))\left(\frac{\lambda z}{1-\overline{\lambda}z}\right)\right]$$

Inserting (9) in the previous equation, it yields:

$$W_V(z) = \frac{\overline{\lambda}\left[p(1-\overline{\lambda}z) + q\lambda z\right]\left[V(z) - V(\overline{\lambda} + \lambda B(z))\right]}{(1-\rho-P_I)(1-\overline{\lambda}z)\left[z - (\overline{\lambda} + \lambda B(z))\right]}\pi_{1,1}. \qquad (38)$$

Which $\pi_{1,1}$ is given by (22). Following Case 2, it gives,

$$W_I(z) \equiv \frac{1}{P_I}\left[\sum_{k=0}^{\infty}\Omega_k\left[B(z)\right]^k\right] = \frac{1}{P_I}G_I(B(z)) = \frac{1}{P_I}\frac{\overline{\lambda}}{\lambda}\left(p\tilde{\pi}_{0,1} + qV(\overline{\lambda} + \lambda B(z))\pi_{1,1}\right). \qquad (39)$$

From Case 3, we have

$$W_B(z)=\frac{1}{\rho}\sum_{k=1}^{\infty}\sum_{i=1}^{\infty}\pi_{k,i}z^{i-1}\left[B(z)\right]^{k-1}==\frac{\bar{\lambda}\left\{p(1-B(z))\tilde{\pi}_{0,1}+\left[1-(p+qB(z))V(\bar{\lambda}+\lambda B(z))\right]\pi_{1,1}\right\}}{\rho\left[\left(\bar{\lambda}+\lambda B(z)\right)-z\right]}. \quad (40)$$

The pgf of waiting time in the queue of a test message when server is on vacation or in idle period is as $W_{OFF}(z)=\frac{1}{(1-\rho)}\left[(1-\rho-P_I)W_V(z)+P_I W_I(z)\right]$. From (38) and (39), we obtain

$$W_{OFF}(z)=\frac{1}{(1-\rho)}\left\{\frac{\bar{\lambda}\left[p(1-\bar{\lambda}z)+q\lambda z\right]\left[V(z)-V(\bar{\lambda}+\lambda B(z))\right]}{(1-\bar{\lambda}z)\left[z-(\bar{\lambda}+\lambda B(z))\right]}\pi_{1,1}+\frac{\bar{\lambda}\left(p\tilde{\pi}_{0,1}+qV(\bar{\lambda}+\lambda B(z))\pi_{1,1}\right)}{\lambda}\right\}. \quad (41)$$

Finally, the pgf of the waiting time in the queue of a test message is given by

$$W_Q(z)=(1-\rho)W_{OFF}(z)+\rho W_B(z)$$

Which implies

$$W_q=\frac{2qV_1+\lambda V_2}{2\left[q+\lambda V_1+pV(\bar{\lambda})\right]}+\frac{\lambda B_2}{2(1-\rho)}. \quad (42)$$

Which is in accordance with $L_{V,p}/\lambda-B_1$ which confirms the result of Little's formula.

OPTIMIZATION ANALYSIS

As a particular case, the Geo/G/1 queueing system with <*T*, *p*> policy, in which the server takes a vacation of fixed length *T* at the ending of the busy period and the server begins service with probability *p* if messages present in the queue at vacation completion instant. We construct the total expected cost function per unit time for the <*T*, *p*> policy system. The main objective of this study is to determine the discrete time T, say T^*, and the probability p, say p^*, simultaneously so that the expected cost function is minimized. To do this, we define the following cost elements:

$C_h \equiv$ Cost per unit time per message present in the system,
$C_s \equiv$ Cost per unit time for a cycle,
$C_r \equiv$ Profit per unit time due to vacation.

Using these cost elements listed in the foregoing and the corresponding system characteristics, the expected cost function $F(T,p)$ per message per unit time is given by

$$F(T,p)=\frac{\lambda}{\left[q+\lambda T+p\bar{\lambda}^T\right]}\left\{C_h\frac{\left[2qT+\lambda T(T-1)\right]}{2}+C_s(1-\rho)\right\}-C_r\frac{\lambda(1-\rho)T}{q+\lambda T+p\bar{\lambda}^T}. \quad (43)$$

The cost function in (43) would have been a hard task to develop analytic results for the optimum value (T^*,p^*) because one is discrete variable *T* and one is continuous variable *p*. We first use direct search method to find the discrete variable, say T^* when p is fixed. Next, we fix T^* and derive the continuous value of p, say p^*.

Direct search method

In practical use, the discrete variable T is bounded by a positive integer T_U. Under a given *p*, we successively use direct substitution of ascendant values of $T=1,2,...,T_U$ into the cost function. The optimum value T^* could be determined by the following:

$$F(T^*|p)=\underset{\rho<1}{Minimize}\,F(T|p),\ T\in\{1,2,...,T_U\} \quad (44)$$

Some numerical examples are presented to demonstrate that the cost function is really convex in T and the solution gives a minimum. For convenience, the numerical experiments are performed by considering $B_1=1.0$ and the following three cases with cost parameter elements: C_h = \$20/message/unit time, C_s = \$1000/unit time, and C_r = \$50/unit time.

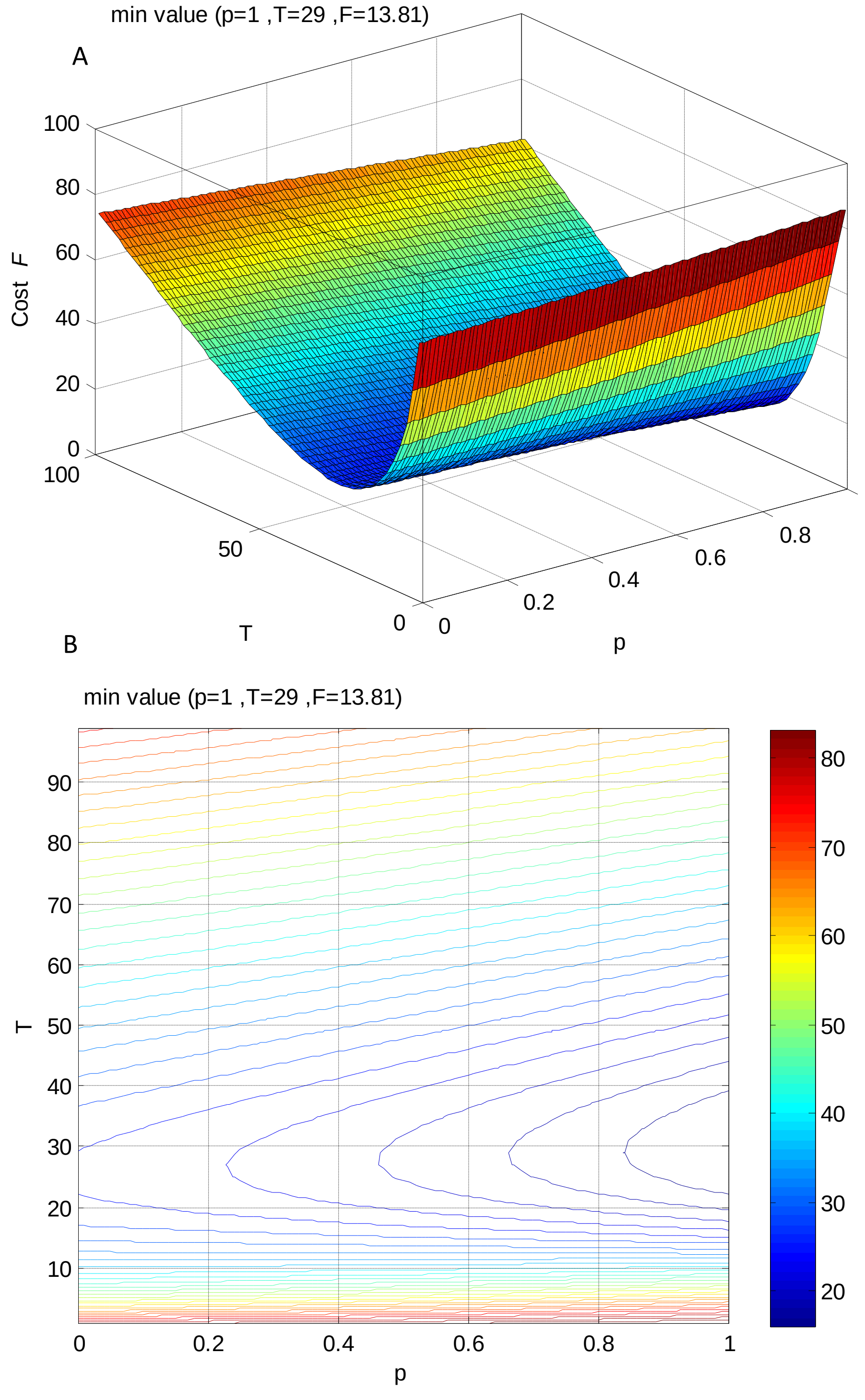

Figure 2. A: Cost with different values of T and p (λ = 0.1, B_1 = 0.1, C_h = \$20, C_s = \$1000, and C_r =\$50). B: Cost contour with different values of T and p (The contour of Figure 2A).

Case 1: λ = 0.1 and vary the values of p and T .
Case 2: λ = 0.5 and vary the values of p and T .
Case 3: λ = 0.9 and vary the values of p and T .

The numerical results are displayed in Figures 2 to 4

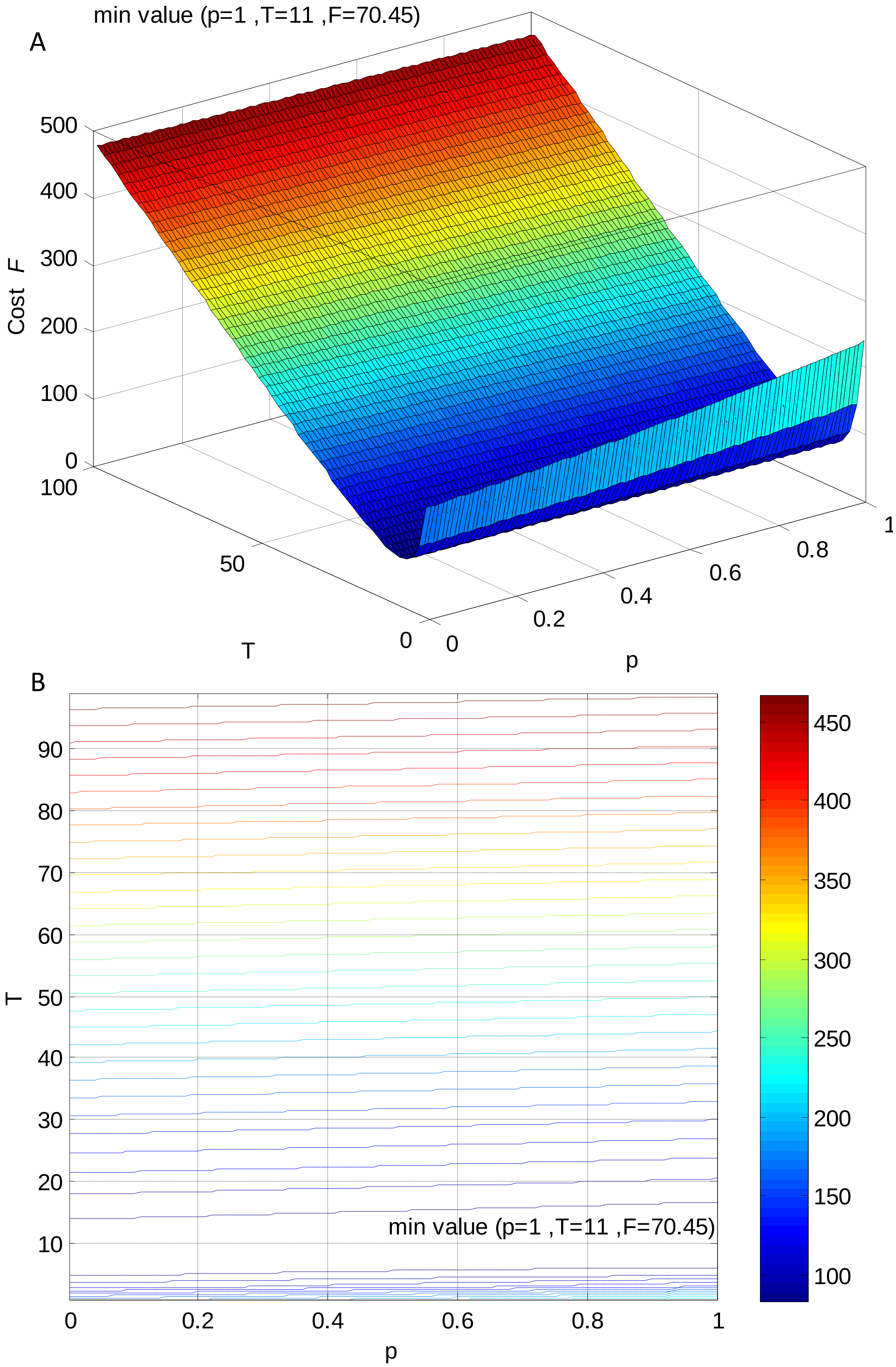

Figure 3. A: Cost with different values of T and p (λ = 0.5, B_1 = 0.1, C_h = \$20, C_s = \$1000, and C_r = \$50). B: Cost contour with different values of T and p (The contour of Figure 3A).

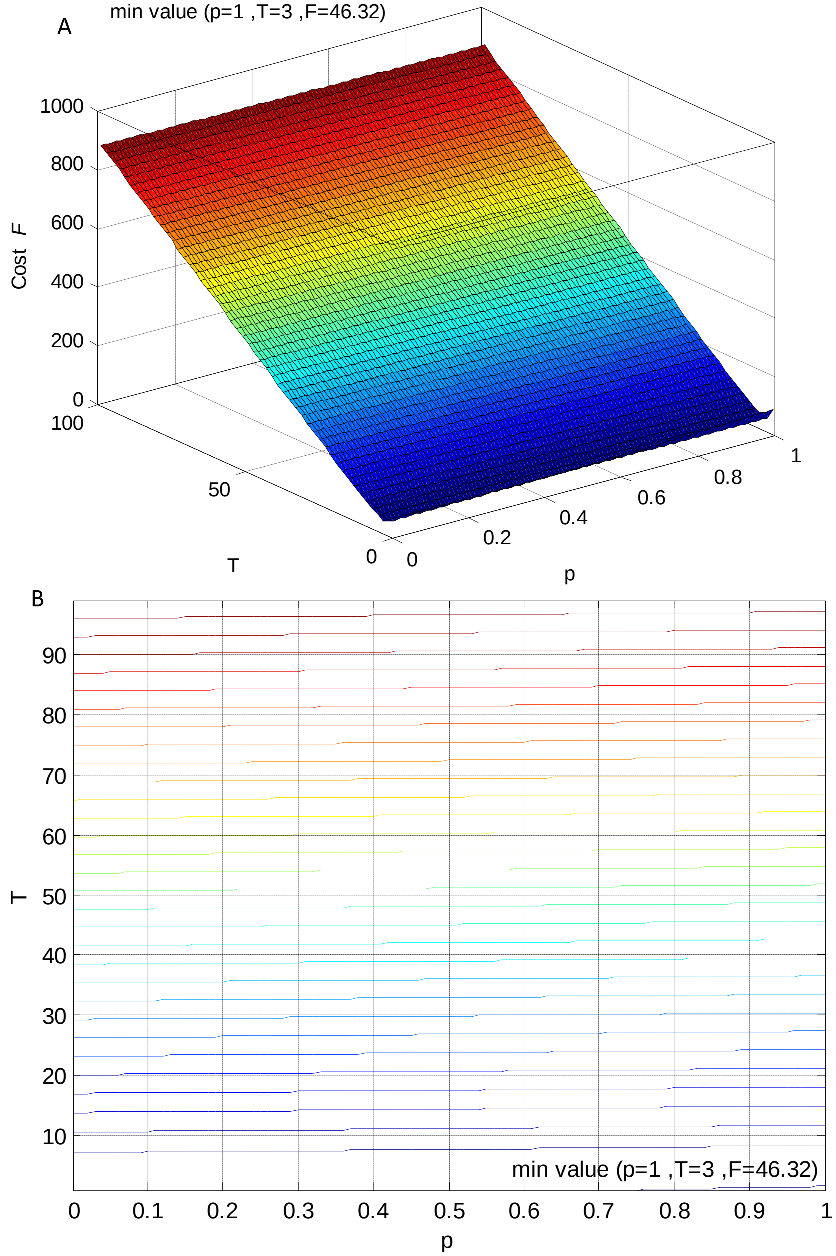

Figure 4. A: Cost with different values of T and p (λ = 0.9, B_1 = 0.1, C_h = \$20, C_s = \$1000, and C_r =\$50). B: Cost contour with different values of T and p (The contour of Figure 4A).

for the three cases, respectively. Figures 2 to 4 (both 3-D and cost contours graphs) show the global optimal values can be obtained.

Optimize p

After we find T^*, we will find p such that the minimum value of $F(T^*, p)$ is achieved, say $F(T^*, p^*)$. The cost minimization problem can be illustrated mathematically as

$$F(T^*, p^*) = \underset{\rho<1}{Minimize}\, F(T^*, p) \tag{45}$$

The study notes that the derivative of the cost function F with respect to p indicates the direction which cost function increases.

Table 1. The illustration of the implement process of p^*.

(λ, B_1)	(0.1, 0.2)	(0.1, 0.5)	(0.1, 0.8)	(0.1, 1.0)	(0.5, 1.0)	(0.9, 1.0)
T^*	31	30	30	30	10	3
p^*	1	1	1	1	1	1
$F^*(T^*, p^*)$	12.46	12.98	13.48	13.8	69.99	46.32
$\frac{dF(p)}{dp}$ in (0,1)	-15.94	-15.63	-15.48	-15.38	-6.02	-2.86
	<0	<0	<0	<0	<0	<0

$$\frac{dF(p)}{dp} = \frac{\lambda}{(q+\lambda T^* + p\overline{\lambda}^{T^*})^2}\Big\{-\frac{1}{2}C_h T^*\left[\lambda T^* + \lambda + 2\overline{\lambda}^{T^*} + \lambda\overline{\lambda}^{T^*}(T^*-1)\right] + (1-\overline{\lambda}^{T^*})(1-\rho)(C_s - C_r T^*)\Big\} \quad (46)$$

The first derivative test supposes that *p* is a critical number on the interval [0, 1] of the continuous cost function F with respect to p. From Equation (46), $\frac{dF(p)}{dp}$ does not change sign at *p*, then F has no local maximum or minimum at *p*.

If

$$C_s < C_r T^* + \frac{1}{2(1-\overline{\lambda}^{T^*})(1-\rho)} C_h T^*\left[\lambda T^* + \lambda + \overline{\lambda}^{T^*} + \lambda\overline{\lambda}^{T^*}(T^*-1)\right]$$

then the derivative $\frac{dF(p)}{dp}$ of equation (46) is negative, and shows that F cost function is decreases on the interval [0, 1] of p. Thus, the cost function F has an absolute minimum (or global minimum) at $p=1$.

If

$$C_s > C_r T^* + \frac{1}{2(1-\overline{\lambda}^{T^*})(1-\rho)} C_h T^*\left[\lambda T^* + \lambda + \overline{\lambda}^{T^*} + \lambda\overline{\lambda}^{T^*}(T^*-1)\right],$$

then the derivative $\frac{dF(p)}{dp}$ of equation (46) is positive on the interval [0, 1] of p, and shows that F cost function increases. Thus, the cost function F has an absolute minimum (or global minimum) at $p=0$.

A numerical illustration is provided by consider the cases:

Case 4: C_h = \$20, C_s = \$1000, C_r = \$50, and vary the values of λ and B_1.

Case 5: λ = 0.5, B_1 = 1.0, C_s = \$1000, C_r = \$20, and vary the values of C_h = 1, 10, 100, 1000 and \$10000.

Case 6: λ = 0.5, B_1 = 1.0, C_h = \$20, C_r = \$20, and vary the values of C_s = 1, 10, 100, 1000 and \$10000.

Case 7: λ = 0.5, B_1 = 1.0, C_h = \$20, C_s = \$1000, and vary the values of C_r = 1, 10, 100, 1000 and \$10000.

For illustrative purpose, we present the four cases listed previously to illustrate the optimization procedure shown in Tables 1 to 4, respectively. The results are in accordance with the analysis listed previously. From Tables 1 to 4, it is seen that (i) T^* increases as λ or B_1 decreases; (ii) T^* increases as C_h decreases or C_s (C_r) increases.

CONCLUSIONS

The study introduces the <*V*, *p*> policy for a discrete-time Geo/G/1 queueing system, in which a single server randomly reactivates when some messages present in the queue at ending of vacation completion instant. Some important system characteristics are derived, including the system length distribution, the turned-off period, the busy period distribution and waiting time distribution.

The study finally develops efficient methods to find the optimal <*T*, *p*> policy that minimizes the expected cost function.

Table 2. The illustration of the implement process of p^*.

C_h	**1**	**10**	**100**	**1000**	**10000**
T^*	45	14	4	1	1
p^*	1	1	1	1	1
$F^*(T^*, p^*)$	12.11	58.21	184.24	245	245
$\frac{dF(p)}{dp}$ in (0,1)	-0.46	-1.68	-13.22	-377.5	-4877.5
	<0	<0	<0	<0	<0

Table 3. The illustration of the implement process of p^*.

C_s	**1**	**10**	**100**	**1000**	**10000**
T^*	1	1	3	10	32
p^*	1	1	1	1	1
$F^*(T^*, p^*)$	-4.75	-2.5	15.38	84.98	301.25
$\frac{dF(p)}{dp}$ in (0,1)	-12.38	-11.25	-10.18	-3.02	-1.17
	<0	<0	<0	<0	<0

Table 4. The illustration of the implement process of p^*.

C_r	**1**	**10**	**100**	**1000**	**10000**
T^*	8	8	9	10	10
p	0	0	0	1	1
$F^*(T^*, p^*)$	543.6	540	462.73	44.99	-4454.13
$\frac{dF(p)}{dp}$ in (0,1)	92.3	91.58	67.6	-11.01	-909.78
	>0	>0	>0	<0	<0

REFERENCES

Bruneel H, Kim BG (1993). Discrete-time models for communication systems including ATM. Kluwer Academic Publishers, Boston.

Doshi BT (1986). Queueing systems with vacations – a survey. Queueing Syst. 1(1):29-66.

Hunter JJ (1983). Mathematical Techniques of Applied Probability, Discrete Time Models: Techniques and Applications, vol. 2. Academic, New York.

Levy Y, Yechiali U (1975). Utilization of idle time in an M/G/1 queueing system. Manage. Sci. 22(2):202-211.

Li J, Tian N (2007).The discrete-time GI/Geo/1 queue with working vacations and vacation interruption. Appl. Math. Comput. 185(1):1-10.

Meisling T (1958). Discrete time queueing theory. Oper. Res. 6(1):96–105.

Takagi H (1991). Queueing Analysis: A Foundation of Performance Evaluation, Vacation and Priority Systems, Part 1, vol. 1. North-

Holland Elsevier, Amsterdam.
Takagi H (1993). Queueing analysis: A foundation of performance evaluation, Discrete-Time Systems, Vol. 3. North-Holland, Amsterdam.
Tian N, Zhang ZG (2002). The discrete-time GI/Geo/1 queue with multiple vacations. Queueing Syst. 40(3):283-294.
Tian N, Zhang ZG (2006). Vacation Queueing Models: Theory and Applications. Springer-Verlag, New York.
Woodward ME (1994). Communication and computer networks: modelling with discrete-time queues. IEEE Computer Society Press, Los Alamitos, CA.
Zhang ZG, Tian N (2001). Geo/G/1 queue with multiple adaptive vacations. Queueing Syst. 38(4):419-429.