

THE OPTIMAL CONTROL IN BATCH ARRIVAL QUEUE WITH SERVER VACATIONS, STARTUP AND BREAKDOWNS

Jau-CHUAN KE*

*National Taichung Institute of Technology
Taichung 404, Taiwan, R.O.C.*

Received: January 2003 / Accepted: January 2004

Abstract: This paper studies the N policy $M^{[x]}/G/1$ queue with server vacations; startup and breakdowns, where arrivals form a compound Poisson process and service times are generally distributed. The server is turned off and takes a vacation whenever the system is empty. If the number of customers waiting in the system at the instant of a vacation completion is less than N , the server will take another vacation. If the server returns from a vacation and finds at least N customers in the system, he is immediately turned on and requires a startup time before providing the service until the system is empty again. It is assumed that the server breaks down according to a Poisson process whose repair time has a general distribution. The system characteristics of such a model are analyzed and the total expected cost function per unit time is developed to determine the optimal threshold of N at a minimum cost.

Keywords: Breakdowns, cost model, queue, startup, vacation.

1. INTRODUCTION

We consider an $M^{[x]}/G/1$ queuing system in which the server operates N policy and is typically a subject to unpredictable breakdowns. The server leaves for a vacation of random length whenever the system becomes empty. At the end of a vacation, the server inspects the system and decides whether to take a vacation, or to take a general startup time. The server immediately starts serving the waiting customers whenever he completes his startup.

* Jau-Chuan Ke. Email: jauchuan@ntit.edu.tw. Postal address: Department of Statistics, National Taichung Institute of Technology, No. 129, Sec. 3, Sanmin Rd., Taichung 404, Taiwan, R.O.C.

Queuing systems with server vacations have attracted much attention from numerous researchers since Levy and Yechiali (1975). Server vacations are useful for the system in which the server wants to utilize his idle time for different purposes. An excellent survey of queuing systems with server vacations was found in Doshi (1986) and Takagi (1991). Queuing systems with the threshold policy (N policy) and multiple vacations, including some applications, were first Lee and Srinivasan (1989) and Kella (1989). They respectively dealt with the batch arrival M/G/1 and the single-unit arrival M/G/1 queuing systems, examined the system performances, and obtained the optimal policy under a stationary cost function. Lee et al. (1994 a, b) further analyzed Lee and Srinivasan's model with a single vacation and multiple vacations, respectively. They provided the probabilistic interpretation of the single (multiple) vacation system with N policy, and their results confirmed the stochastic decomposition property given by Fuhrmann and Cooper (1985).

The server startup corresponds to the preparatory work of the server before starting the service. In some actual situations, the server often requires a startup time before starting his each service period. Examining queuing systems which combine the N policy with startup time, Baker (1973) first proposed the N policy M/M/1 queuing system with exponential startup time. Borthakur et al. (1987) extended Baker's results to the general startup time. The N policy M/G/1 queuing system with startup time was first studied by Minh (1988) and was investigated by several researchers such as Medhi and Templeton (1992), Takagi (1993), Lee and Park (1997), and others. Recently, Hur and Paik (1999) examined the operation characteristics of the N policy M/G/1 queuing system with server startup and explained how the system's optimal policy and cost behave for various arrival rates.

Most studies on queuing systems use "perfect" (reliable) servers. In many real systems, however, the server may meet unpredictable breakdowns. Therefore, queuing models with server breakdowns are more realistic representation of the systems. Discussing queuing systems with N policy and server breakdowns, Wang (1995) proposed the N policy M/M/1 queuing system with server breakdowns at first. Wang (1997) and Wang et al. (1999) extended Wang's model to the N policy M/E_k/1 and M/H₂/1 queuing system cases, respectively. They developed the analytic closed-form solutions and provided a sensitivity analysis. Comparable work on the batch arrival queues where the server operates an N policy with vacations, startup and breakdowns is rarely found in the literature. This motivates us to develop the N policy for an M^[x]/G/1 queuing system where the server is characterized with vacations, startup and breakdowns.

In this paper, we develop the probability-generating function of the number of customers present when the server begins performing startup. We also derive other important system characteristics, such as the expected number of customers in the system, waiting time distribution in the queue, the expected length of the idle period and busy period, etc. Following the construction of the expected cost function per unit time, an efficient and quick procedure is developed for searching the optimal threshold N^* value that minimize the cost function. Some numerical examples are also presented.

2. THE SYSTEM ASSUMPTIONS AND NOTATIONS

In this paper, we consider the control policy for an $M^{[x]}/G/1$ queuing system with the following specifications. It is assumed that customers arrive in batches to occur according to a compound Poisson with rate λ . Let X_k denote the number of customers belonging to the k th arrival batch, where X_k , $k = 1, 2, 3, \dots$, are with a common distribution, where

$$\Pr[X_k = n] = g_n, \quad n = 1, 2, 3, \dots$$

The service times for a customer are independent and identically distributed random variables (S) with a common distribution function $S(t)$. The server is subject to breakdowns at any time with Poisson breakdown rate α when he is working. Whenever the server fails, it is immediately repaired at a repair facility, where the repair times are independent and identically distributed random variables (R) with a common distribution function $R(t)$. If at any time any customer arrives, he goes to the service facility for service. Arriving customers 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 customer at a time. 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 a broken down server, customers continue to arrive according to a Poisson process.

Furthermore, we consider the system in which the server is turned off and takes a sequence of vacations for a period of random length V whenever the system becomes empty. When the server returns from a vacation and finds that the number of customers waiting in the system is less than N , he will go on a vacation again. If N or more customers are accumulated in the system, the server is immediately turned on but is temporarily unavailable to the waiting customers. He needs a startup time with random length U before starting his service. As soon as the server finishes startup, he starts 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 repaired. Once the broken down server is repaired, he immediately returns to serve a customer until the system is empty.

For any discrete random variable, Y , used in the analysis, we adopt the following notation:

$$Y(z) = \sum_{i=1}^{\infty} z^i y_i,$$

the probability generating function (p.g.f.) of Y , where $y_i = \Pr[Y = i]$.

Similarly, for any continuous random variable Z used in the analysis, $Z(t)$ denotes the distribution function of Z , $Z^*(\theta) = \int_0^{\infty} e^{-\theta t} dZ(t)$ denotes its Laplace-Stieltjes transform (LST) of and $Z^{*(l)}(\theta)$ the l th order derivative of $Z^*(\theta)$ with respect to θ .

3. PROBABILITY GENERATING FUNCTION

In this section we first construct the probability generating function of the number of customers present when the server begins performing startup. Then, we study various system characteristics by means of the probability generating function.

Let us define that B is the number of batches that arrive during a vacation. Using the definition of Poisson arrivals, $e^{-\lambda t} (\lambda t)^k / k!$ is the probability that k batches arrive during $[0, t]$. We have

$$b_k = \Pr(B = k) = \int_0^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} dV(t).$$

For convenience of computation of b_k , we use the following equivalent expressions:

$$b_k = \frac{(-\lambda)^k}{k!} V^{*(k)}(\lambda).$$

Furthermore, we define that Q is the number of customers that arrive during a vacation. Using the property of convolution, the probability that there are exact n customers during k batches arriving as follow:

$$g_n^{(k)} = \overbrace{g_n \otimes g_n \otimes \cdots \otimes g_n}^{k \text{ } g_n \text{'s}} = \Pr[X_1 + X_2 + \cdots + X_k = n],$$

where $g_0^{(0)}$ is defined to be 1.

From the definitions of Q and $g_n^{(k)}$, we have the following relations

$$q_n = \Pr(Q = n) = \sum_{k=0}^n b_k g_n^{(k)}, \quad n = 0, 1, 2, \dots \quad (1)$$

From (1), we obtain the following *Theorem 1*.

Theorem 1. *The probability generating function of Q is given by*

$$Q(z) = V^*(\lambda - \lambda X(z)).$$

Differentiating $Q(z)$ with respect to z , it finally yields the first and second factorial moments of Q as follows:

$$Q^{(1)} = E[Q] = \lambda E[X]E[V], \quad (2a)$$

and

$$Q^{(2)} = E[Q(Q-1)] = \lambda E[X(X-1)]E[V] + (\lambda E[X])^2 E[V^2]. \quad (2b)$$

Let Π_N be the number of customers present when *the server begins performing startup*. The following result, stated as *Theorem 2* expresses the p.g.f. of Π_N .

Theorem 2.

$$\Pi_N(z) = \frac{1}{1-q_0} \left\{ \sum_{j=1}^{N-1} q_j z^j [\Pi_{N-j}(z) - 1] + \sum_{k=1}^{\infty} z^k q_k \right\}, \quad N = 1, 2, 3, \dots \quad (3)$$

Proof: Conditioning on the number of customers who arrive during the previous vacation, we have

$$\Pr[\Pi_N = k] = \sum_{j=0}^{N-1} q_j \Pr[\Pi_{N-j} = k - j] + q_k, \quad k \geq N, \quad N = 1, 2, 3, \dots \quad (4)$$

It is easy to solve (4) in term of $\Pr[\Pi_N = k]$, and it yields

$$\Pr[\Pi_N = k] = \frac{1}{1-q_0} \left\{ \sum_{j=1}^{N-1} q_j \Pr[\Pi_{N-j} = k - j] + q_k \right\}, \quad k \geq N, \quad N = 1, 2, 3, \dots \quad (5)$$

From (5), we get the p.g.f. of Π_N is obtained as

$$\Pi_N(z) = \frac{1}{1-q_0} \left\{ \sum_{k=N}^{\infty} z^k \sum_{j=1}^{N-1} q_j \Pr[\Pi_{N-j} = k - j] + \sum_{k=N}^{\infty} z^k q_k \right\}. \quad (6)$$

Simplify (6) to obtain the desired result of (3). ■

Using (3), $E[\Pi_N]$ and $E[\Pi_N(\Pi_N - 1)]$ are obtained by taking first and second order derivative of $\Pi_N(z)$ with respect to z .

Let $\pi_N^{(1)} = E[\Pi_N]$ and $\pi_N^{(2)} = E[\Pi_N(\Pi_N - 1)]$. Thus we have as follows:

$$\pi_N^{(1)} = \frac{1}{1-q_0} \left\{ \sum_{j=1}^{N-1} q_j \pi_{N-j}^{(1)} + Q^{(1)} \right\}, \quad N = 1, 2, 3, \dots \quad (7)$$

and

$$\pi_N^{(2)} = \frac{1}{1-q_0} \left\{ \sum_{j=1}^{N-1} q_j [2j\pi_{N-j}^{(1)} + \pi_{N-j}^{(2)}] + Q^{(2)} \right\}, \quad N = 1, 2, 3, \dots \quad (8)$$

where \sum_a^b indicates zero when $a > b$, and $Q^{(1)}$ and $Q^{(2)}$ are given in (2).

To complete the computations of $\pi_N^{(1)}$ and $\pi_N^{(2)}$, we may obtain $\pi_k^{(1)}$ and $\pi_k^{(2)}$ ($1 \leq k \leq N$) recursively using (7) and (8).

4. SYSTEM CHARACTERISTICS

4.1. Expected number of the customers in the system

We recall some results in the ordinary $M^{[x]}/G/1$ queuing system with server breakdowns. Let H be a random variable representing the completion time of a customer service, which includes both the service time of a customer and the repair time of a server. The useful results of Tang (1997) are as follows;

$$E[H] = E[S](1 + \alpha E[R]), \quad (9)$$

$$E[H^2] = (1 + \alpha E[R])^2 (E[S^2]) + \alpha E[S]E[R^2], \quad (10)$$

and

$$\rho_H = \lambda E[X]E[H] = \rho(1 + \alpha E[R]). \quad (11)$$

where $\rho = \lambda E[X]E[S]$. Note that ρ_H is traffic intensity and it should be assumed to be less than unity.

Applying the results of Chae and Lee (1995) and Tang (1997), we obtain the expected waiting time in the queue for the N policy $M^{[x]}/G/1$ queuing system with server vacations, startup and breakdowns given by

$$W_q = (L_q + E[X_r])E[H] + \rho_H E[H_r] + (1 - \rho_H)E[I_r], \quad (12)$$

where L_q denotes the expected number of customers in the queue, X_r denotes the residual batch size, H_r denotes the residual completion time and I_r denotes the residual idle time.

Substituting $\rho_H W_q$ for $L_q E[H]$ in (12), and then solving for W_q yields

$$W_q = \frac{E[H]E[X_r]}{1 - \rho_H} + \frac{\rho_H E[H_r]}{1 - \rho_H} + E[I_r]. \quad (13)$$

where ρ_H is given in (11).

From Chae and Lee (1995), we obtain $E[X_r]$ given by

$$E[X_r] = \frac{1}{2} \left(\frac{E[X^2]}{E[X]} - 1 \right). \quad (14)$$

It is well-known results of the Kleinrock (1975) that $E[H_r]$ is given by

$$E[H_r] = \frac{E[H^2]}{2E[H]}. \quad (15)$$

Let V_r and U_r denote the residual vacation time and the residual startup time. Following the result of Chae and Lee (1995) and using the definition of Π_N , the expected residual vacation time is given by

$$E[V_r] = \frac{1}{\lambda E[X]} \left\{ \frac{1}{2} \left(\frac{E[\Pi_N^2]}{E[\Pi_N]} - 1 \right) - E[X_r] \right\}. \quad (16)$$

Inserting (14) in (16), we obtain

$$E[V_r] = \frac{1}{2\lambda E[X]} \left(\frac{\pi_N^{(2)}}{\pi_N^{(1)}} - \frac{E[X(X-1)]}{E[X]} \right). \quad (17)$$

According to the results of Takagi (1991, section 2.2) and Chae and Lee (1995), we have

$$E[U_r] = \frac{2E[U] + \lambda E[U^2]}{2(1 + \lambda E[U])}. \quad (18)$$

Since the idle time is consisted of the vacation time and the startup time, we have the expected residual idle time as follows:

$$E[I_r] = E[V_r] + E[U_r].$$

From (17) and (18), we have

$$E[I_r] = \frac{1}{2\lambda E[X]} \left(\frac{\pi_N^{(2)}}{\pi_N^{(1)}} - \frac{E[X(X-1)]}{E[X]} \right) + \frac{2E[U] + \lambda E[U^2]}{2(1 + \lambda E[U])}. \quad (19)$$

Finally follows from (9)-(10), (13)-(15) and (19), after using Little's rule ($L_{vs} = L_q + \rho_H = \lambda W_q + \rho_H$) we obtain an expression for L_{vs} , stated as *Theorem 3*.

Theorem 3. *The expected number of customers for the N policy $M^{[x]}/G/1$ queuing system server vacations, startup and breakdowns is given by*

$$\begin{aligned} L_{vs} = & \frac{\lambda^2 E[X](1 + \alpha E[R])^2 E[S^2] + \lambda \alpha \rho E[R^2]}{2(1 - \rho_H)} \\ & + \frac{\lambda E[S](1 + \alpha E[R])E[X(X-1)]}{2(1 - \rho_H)E[X]} \\ & + \frac{1}{2E[X]} \left(\frac{\pi_N^{(2)}}{\pi_N^{(1)}} - \frac{E[X(X-1)]}{E[X]} \right) \\ & + \frac{2\lambda E[U] + \lambda^2 E[U^2]}{2(1 + \lambda E[U])} + \rho_H. \end{aligned} \quad (20)$$

4.2. Expected length of the complete period, the idle period and the busy cycle

The server remains on vacations until he finds N or more customers after returning from a vacation. This is called the vacation period. The startup period begins when the server performs startup as soon as the server returns from a vacation and finds at least N customers. This startup period terminates when he starts providing the service.

The customers arriving during startup and vacation periods will not be immediately served until the server completes the startup. Therefore, the idle period is the sum of the vacation and startup period. The busy period is initiated when the server completes his startup and starts serving the waiting customers. During the busy period, the server may break down and start his repair immediately. This is called the breakdown period. After the server is repaired, he returns and provides service until there are no customers in the system. Since the complete period starts when the startup period is over and terminates when there are no customers in the system, the complete period is represented by the sum of the busy period and the breakdown period.

First let $H_o^*(\theta)$ denote the LST of the complete period in the ordinary $M^{[x]}/G/1$ queuing system with server breakdowns and $H_{os}^*(\theta)$ be the LST of the complete period in ordinary the $M^{[x]}/G/1$ queuing system with server startup time. It is useful results of Baba (1986), Tang (1997), and Wang (1997) that the LST of the complete period started with one customer in the $M/G/1$ queuing system with server breakdowns can be expressed by

$$H_o^*(\theta) = H^*[\theta + \lambda - \lambda X(H_o^*(\theta))], \quad (22)$$

where $H^*(\cdot)$ is LST of completion time H .

Applying the well-known results of Takagi (1991, section 2.2) and Tang (1997) we have the following important results

$$H_{os}^*(\theta) = H_o^*(\theta)U^*[\lambda - \lambda X(H_o^*(\theta))], \quad (23)$$

where $H_o^*(\theta)$ is given in (22).

Differentiating (22) with respect to θ , we derive the expected length of the complete period initiated with one customer in the system as

$$E[H_o] = \frac{E[H]}{1 - \rho_H}. \quad (24)$$

where $E[H]$ and ρ_H are given in (9) and (11), respectively.

Differentiate (23) with respect to θ and using (24), it follows that

$$\begin{aligned} E[H_{os}] &= E[H_o] + \lambda E[X]E[U]E[H_o] \\ &= \frac{E[H]}{1 - \rho(1 + \alpha E[R])} + \frac{\lambda E[X]E[U]E[H]}{1 - \rho(1 + \alpha E[R])}. \end{aligned} \quad (25)$$

Expected length of the complete period

Let $H_{vs}^*(\theta)$ be the LST of complete period for the N policy $M^{[x]}/G/1$ queuing system with server vacations, startup and breakdowns, and the first customer in each complete period needs to wait for a server startup time U before receiving the service. After the server finishes startup, he resumes the supply of service. The time spent on serving the customers who arrive during the vacation period and the subsequent arrivals

while the server is providing service is denoted by H_v with LST $H_v^*(\theta)$. And H_s denotes the time spent on serving the arrivals during the server startup period and those subsequent arrivals before the system becomes empty again, where $H_s^*(\theta)$ is the LST of H_s . It is clear that $H_{vs}^*(\theta)$ is convolution of $H_v^*(\theta)$ and $H_s^*(\theta)$. Since the vacation and the startup times are independent, we have

$$H_{vs}^*(\theta) = H_v^*(\theta) \otimes H_s^*(\theta) = H_v^*(\theta)H_s^*(\theta). \quad (26)$$

Find $E[H_s]$

Recall that the expected length of complete period given by (25) decomposes into two parts: one (the first term) is the expected length of complete period in the ordinary $M^{[x]}/G/1$ queuing system with server breakdowns, the other (the second term) is the expected length of the complete period caused by the server startup. Using the meanings of the above listed H_s we have

$$E[H_s] = \frac{\lambda E[X]E[U]E[H]}{1 - \rho(1 + \alpha E[R])} = \frac{\rho_H E[U]}{1 - \rho_H}. \quad (27)$$

Find $E[H_v]$

Since $H_o^*(\theta)$ is the LST of the complete period started with one customer and arrival process is Poisson, the LST of the complete period initiated with n customers in the system is given by $[H_o^*(\theta)]^n$. Thus, using the definitions of H_v we have

$$H_v^*(\theta) = \sum_{k=N}^{\infty} \Pr[\Pi_N = k] (H_o^*(\theta))^k = \Pi_N(H_o^*(\theta)). \quad (28)$$

Differentiating (28) with respect to θ and using (24), we get

$$E[H_v] = \frac{\pi_N^{(1)} E[H]}{1 - \rho(1 + \alpha E[R])} = \frac{\pi_N^{(1)} E[S](1 + \alpha E[R])}{1 - \rho(1 + \alpha E[R])}. \quad (29)$$

Differentiating (26) with respect to θ and using (27) and (29), we obtain an expression for $E[H_{vs}]$, stated as *Theorem 4*.

Theorem 4. *The expected length of the complete period for the N policy $M^{[x]}/G/1$ queuing system with server vacations, startup and breakdowns is given by*

$$E[H_{vs}] = (\pi_N^{(1)} + \lambda E[X]E[U]) \frac{E[S](1 + \alpha E[R])}{1 - \rho(1 + \alpha E[R])}. \quad (30)$$

Expected length of the busy period and the breakdown period

The expected length of the busy period and the expected length of breakdown period are denoted by $E[B_{vs}]$ and $E[D_{vs}]$, respectively. Recall that the complete period is the sum of the busy and the breakdown periods which imply $E[H_{vs}] = E[B_{vs}] + E[D_{vs}]$. Hence, from (30) we have

$$E[B_{vs}] = \frac{(\pi_N^{(1)} + \lambda E[X]E[U])E[S]}{1 - \rho(1 + \alpha E[R])}, \quad (31)$$

and

$$E[D_{vs}] = \frac{(\pi_N^{(1)} + \lambda E[X]E[U]) \alpha E[S]E[R]}{1 - \rho(1 + \alpha E[R])}. \quad (32)$$

Expected length of the idle period

Let $I_{vs}^*(\theta)$ be the LST of the idle period, which is the sum of vacation period and startup period, for the N policy $M^{[x]}/G/1$ queuing system with server vacations, startup and breakdowns. Suppose $I_v^*(\theta)$ denotes the LST of the vacation period. It is clear that $I_{vs}^*(\theta)$ is convolution of $I_v^*(\theta)$ and $U^*(\theta)$. Since the vacation time and the startup time are independent, we have

$$I_{vs}^*(\theta) = I_v^*(\theta) \otimes U^*(\theta) = I_v^*(\theta)U^*(\theta). \quad (33)$$

Since $\pi_N^{(1)}$ is the expected number of arrivals before the server starts performing startup, we have the expected length of the vacation period as follows;

$$E[I_v] = \frac{\pi_N^{(1)}}{\lambda E[X]}. \quad (34)$$

Differentiating (33) with respect to θ and using (34), the expected length of the idle period for this system is expressed by the following theorem.

Theorem 5.

$$E[I_{vs}] = \frac{\pi_N^{(1)}}{\lambda E[X]} + E[U]. \quad (35)$$

Expected length of the busy cycle

The busy cycle for the N policy $M^{[x]}/G/1$ queuing system with server vacations, startup and breakdowns, denoted by Θ_{vs} , is the length of time from the beginning of the last idle period to the beginning of the next idle period. From (30) and (35), we have

$$E[\Theta_{vs}] = \frac{\pi_N^{(1)} + \lambda E[X]E[U]}{\lambda E[X][1 - \rho(1 + \alpha E[R])]} \quad (36)$$

5. SPECIAL CASES

In this section, we present some existing results in the literature which are special cases of our system.

Case 1: It is to be noted that when $\Pr[U = 0] = 1$, $\alpha = 0$ and $N = 1$, we recover the results for the ordinary $M^{[x]}/G/1$ queuing system with multiple vacations (i.e. a sequence of the same vacations of random length). In this case, the expected residual vacation time, $E[V_r]$ given in equation (17), can be reduced to a special case for the last term of expression (3.21a) of Takagi (1991, p143) or the expression (6) of Chae and Lee (1995).

Case 2: When $\Pr[U = 0] = 1$, $\alpha = 0$ and $N > 1$, we recover the results for the N policy $M^{[x]}/G/1$ queuing system with multiple vacations. In this case, expression (7) for $\pi_N^{(1)}$, expression (8) for $\pi_N^{(2)}$, and (36) for $E[\Theta_{vs}]$ can be reduced to a special case for the expressions (4.10), (4.11) and (4.38) of Lee and Srinivasan (1989).

6. OPTIMAL POLICY

In this section, we develop the total expected cost function per unit time for the N policy $M^{[x]}/G/1$ queuing system with server vacations, startup, and breakdowns, in which N is decision variable. Let us consider the following costs:

- c_s \equiv setup cost for per busy cycle;
- c_h \equiv holding cost per unit time for each customer present in the system;
- c_r \equiv startup cost per unit time for the preparatory work of the server before starting the service;
- c_o \equiv cost per unit time for keeping the server on and in operation;
- c_d \equiv breakdown cost per unit time for a failed server.

Using the definitions of each cost element and its corresponding system performances, the total expected cost with threshold N becomes

$$F(N) = \frac{c_s}{E[\Theta_{vs}]} + c_h L_{vs} + c_r \frac{E[U]}{E[\Theta_{vs}]} + c_o \frac{E[B_{vs}]}{E[\Theta_{vs}]} + c_d \frac{E[D_{vs}]}{E[\Theta_{vs}]} \quad (37)$$

Substituting (21), (31)-(32), and (36) into (37), we obtain

$$\begin{aligned}
F(N) &= \frac{\lambda E[X](1-\rho_H)(c_s + c_r E[U])}{\pi_N^{(1)} + \lambda E[X]E[U]} \\
&+ c_h \left[\frac{\lambda^2 E[X](1+\alpha E[R])^2 E[S^2] + \lambda \alpha \rho E[R^2]}{2(1-\rho_H)} \right. \\
&+ \frac{\lambda E[S](1+\alpha E[R])E[X(X-1)]}{2(1-\rho_H)E[X]} + \frac{1}{2E[X]} \left(\frac{\pi_N^{(2)}}{\pi_N^{(1)}} - \frac{E[X(X-1)]}{E[X]} \right) \\
&+ \left. \frac{2\lambda E[U] + \lambda^2 E[U^2]}{2(1+\lambda E[U])} + \rho_H \right] + c_d \rho \alpha E[R] + c_o \rho \\
&= \frac{A_1}{\pi_N^{(1)} + \lambda E[X]E[U]} + \frac{c_h}{2E[X]} \frac{\pi_N^{(2)}}{\pi_N^{(1)}} + A_2,
\end{aligned} \tag{38}$$

where $A_1 = \lambda E[X](1-\rho_H)(c_s + c_r E[U])$,

$$\begin{aligned}
A_2 &= c_h \left[\frac{\lambda^2 E[X](1+\alpha E[R])^2 E[S^2] + \lambda \alpha \rho E[R^2]}{2(1-\rho_H)} \right. \\
&+ \frac{\lambda E[S](1+\alpha E[R])E[X(X-1)]}{2(1-\rho_H)E[X]} + \frac{2\lambda E[U] + \lambda^2 E[U^2]}{2(1+\lambda E[U])} \\
&\left. - \frac{E[X(X-1)]}{2(E[X])^2} + \rho_H \right] + c_d \rho \alpha E[R] + c_o \rho,
\end{aligned}$$

and ρ_H is given by (11).

We want to find the value of N that yields the minimum expected cost per unit of time. Following the concepts of Lee and Srinivasan (1989), the optimal value N^* of N is given by the first N such that $I(N) > 0$. That is

$$N^* = \min \{ N \geq 1 \mid I(N) > 0 \}. \tag{39}$$

where

$$\begin{aligned}
I(N) &= F(N+1) - F(N) \\
&= A_1 \left[\frac{1}{\pi_{N+1}^{(1)} + \lambda E[X]E[U]} - \frac{1}{\pi_N^{(1)} + \lambda E[X]E[U]} \right] \\
&+ \frac{c_h}{2E[X]} \left[\frac{\pi_{N+1}^{(2)}}{\pi_{N+1}^{(1)}} - \frac{\pi_N^{(2)}}{\pi_N^{(1)}} \right].
\end{aligned} \tag{40}$$

The sign of $I(N)$ determines whether $F(N)$ increases or decreases with respect to N . Numerical experiments based on (39) can convince us that the expected cost function is convex which is performed in the following section.

7. NUMERICAL EXAMPLES

We present some numerical examples in this section to demonstrate how the management of a production line system uses the above results to make the decision to minimize the total expected cost function regarding the threshold value.

Consider a production line in which the production does not start until at least some specified number of units, N ($N \geq 1$), are accumulated during an idle period. The units arrive in batches according to a compound Poisson process with batch arrival rate $\lambda = 0.6$. The operator will take a sequence vacations whenever there are no units to process. The operator of the machine may utilize his vacation time to perform some extra operations such as preventive maintenance, or some other work, etc. When the operator returns from a vacation and finds that the number of processed units is less than threshold, N , he will go on a vacation again. It is noted that the operator requires a startup time to operate machine when he finds that the number of processed units reaches or exceeds N and the operation may be interrupted because of emergent events. We wish to obtain an optimal policy regarding the threshold value, which minimizes the total expected cost function and the costs are assumed as follows;

$$c_s = 1000, c_h = 1, c_r = 100, c_d = 200, c_o = 100.$$

The above system can be modeled as $M^{[x]}/G/1$ queue with the following assumptions:

- Batch arrival size distribution is geometric with parameter p ;
- vacation time is a 2-stage Erlang distributions with mean $E[V]$;
- startup time is an exponential distribution with mean $E[U]$;
- service time is a 2-type hyperexponential distribution with probability density function $f(s) = \xi_1 \mu_1 \exp(-\mu_1 s) + \xi_2 \mu_2 \exp(-\mu_2 s)$ and mean $E[S]$ (We let $\xi_1 = 3/4, \xi_2 = 1/4, \mu_1 = 3,$ and $\mu_2 = 1$ in all numerical examples, and therefore $E[S] = 0.5$);
- repair time of the server is a 3-stage Erlang distribution with mean $E[R]$.

First, we select $\lambda = 0.6, p = 0.55, E[S] = 0.5, E[V] = 5.0, E[U] = 1.0,$ and change in specific values $(\alpha, E[R])$. The numerical results of the optimal threshold and its minimum expected cost are summarized in Table 1. One easily sees from Table 1 that the optimal threshold and its minimum expected cost slightly change when α is from 0.05 to 0.3 or $E[R]$ from 0.2 to 1.0.

Table 1: The optimal threshold N^* and its minimum expected cost $F(N^*)$

$$(\lambda = 0.6, p = 0.55, E[S] = 0.5, E[V] = 5.0, E[U] = 1.0).$$

$(\alpha, E[R])$	(0.2, 0.2)	(0.2, 0.5)	(0.2, 1)	(0.3, 0.5)	(0.1, 0.5)	(0.05, 0.5)
N^*	25	25	24	24	25	25
$F(N^*)$	59.982	62.716	67.263	64.981	60.439	59.301

Finally, we select $\lambda = 0.6$, $p = 0.45$, $\alpha = 0.05$, $E[S] = 0.5$, $E[R] = 0.4$, and change in specific values $(E[V], E[U])$. The numerical results of the optimal threshold and its minimum expected cost are summarized in Table 2. We observe from Table 2 that the optimal threshold N^* and the minimum expected cost $F(N^*)$ increase as $E[V]$ or $E[U]$ increase.

Table 2: The optimal threshold N^* and its minimum expected cost $F(N^*)$
($\lambda = 0.6$, $p = 0.45$, $\alpha = 0.05$, $E[S] = 0.5$, $E[R] = 0.4$).

$(E[V], E[U])$	(10, 2)	(5, 2)	(1, 2)	(5, 5)	(5, 10)	(5, 15)
N^*	35	34	12	36	39	41
$F(N^*)$	74.139	70.423	50.322	74.743	81.237	87.130

8. CONCLUSIONS

For the N policy $M^{[x]}/G/1$ queuing system, we easily derived some important system characteristics through the key probability - generating function and decomposition property. We also performed a numerical analysis between the optimal value N and the specific values of system parameters. The model is significant since more general situations in practical applications are considered in the model.

REFERENCES

- [1] Baba, Y., "On the $M^x/G/1$ queue with vacation time", *Operations Research Letters*, 5 (1986) 93-98.
- [2] Baker, K.R., "A note on operating policies for the queue $M/M/1$ with exponential startup", *INFOR*, 11 (1973) 71-72.
- [3] Borthakur, A., Medhi, J., and Gohain, R., "Poisson input queueing systems with startup time and under control operating policy", *Computers Ops Res.*, 14 (1987) 33-40.
- [4] Chae, K.C., and Lee, H.W., " $M^x/G/1$ vacation models with N -policy: Heuristic interpretation of the mean waiting time", *Journal of the Operational Research Society*, 46 (1995) 258-264.
- [5] Doshi, B.T., "Queueing system with vacations – A survey", *Queueing Systems*, 1 (1986) 29-66.
- [6] Fuhrmann, S.W., and Cooper, R.B., "Stochastic decompositions in the $M/G/1$ queue with generalized vacations", *Operations Research*, 33 (1985) 1117-1129.
- [7] Hur, S., and Paik, S.J., "The effect of different arrival rates on the N -policy of $M/G/1$ with server setup", *Applied Mathematical Modelling*, 23 (1999) 289-299.
- [8] Kella, O., "The threshold policy in the $M/G/1$ queue with server vacations", *Naval. Res. Logist.*, 36 (1989) 111-123.
- [9] Kleinrock, L., *Queueing Systems, Vol. 1*, Wiley, New York, 1975.
- [10] Lee, S.S., Lee, H.W., and Chae, K.C., "On a batch arrival queue with N policy and single vacation", *Computers Ops Res.*, 22 (1995) 173-189.

- [11] Lee, H.W., Lee, S.S., Park, J.O., and Chae, K.C., "Analysis of $M^x/G/1$ queue with N policy and multiple vacations", *J. Appl. Prob.*, 31 (1994) 467-496.
- [12] Lee, H.W., and Park, J.O., "Optimal strategy in N -policy production system with early set-up", *Journal of the Operational Research Society*, 48 (1997) 306-313.
- [13] Lee, H.S., and Srinivasan, M.M., "Control policies for the $M^x/G/1$ queueing system", *Management Science*, 35 (1989) 708-721.
- [14] Levy, Y., and Yechiali, U., "Utilization of idle time in an $M/G/1$ queueing system", *Management Science*, 22 (1975) 202-211.
- [15] Medhi, J., and Templeton, J.G.C., "A Poisson input queue under N -policy and with a general start up time", *Computers Opns Res.*, 19 (1992) 35-41.
- [16] Minh, D.L., "Transient solutions for some exhaustive $M/G/1$ queues with generalized independent vacations", *European Journal of Operational Research*, 36 (1988) 197-201.
- [17] Takagi, H., *Queueing Analysis: A Foundation of Performance Evaluation, Vol. I, Vacation and Priority Systems, Part I*, North-Holland, Amsterdam, 1991.
- [18] Takagi, H., " $M/G/1/K$ queues with N -policy and setup times", *Queueing Systems*, 14 (1993) 79-98.
- [19] Tang, Y.H., "A single-server $M/G/1$ queueing system subject to breakdowns - some reliability and queueing problem", *Microelectron. Reliab.*, 37 (1997) 315-321.
- [20] Wang, K.-H., "Optimal operation of a Markovian queueing system with a removable and non-reliable server", *Microelectron. Reliab.*, 35 (1995) 1131-1136.
- [21] Wang, K.-H., "Optimal control of an $M/E_k/1$ queueing system with removable service station subject to breakdowns", *Journal of the Operational Research Society*, 48 (1997) 936-942.
- [22] Wang, K.-H., Chang, K.-W., and Sivazlian, B.D., "Optimal control of a removable and non-reliable server in an infinite and a finite $M/H_2/1$ queueing system", *Applied Mathematical Modelling*, 23 (1999) 651-666.