Comparative analysis of a randomized $N$-policy queue: An improved maximum entropy method

Kuo-Hsiung Wang $^{a,}$*, Dong-Yuh Yang $^{b}$, W.L. Pearn $^{c}$

$^{a}$ Department of Business Administration, Asia University, Wufeng, Taichung 41354, Taiwan
$^{b}$ Institute of Information Science and Management, National Taipei College of Business, Taipei 100, Taiwan
$^{c}$ Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu 30050, Taiwan

**Abstract**

We analyze a single removable and unreliable server in an M/G/1 queueing system operating under the $(p,N)$-policy. As soon as the system size is greater than $N$, turn the server on with probability $p$ and leave the server off with probability $(1 - p)$. All arriving customers demand the first essential service, where only some of them demand the second optional service. He needs a startup time before providing first essential service until there are no customers in the system. The server is subject to break down according to a Poisson process and his repair time obeys a general distribution. In this queueing system, the steady-state probabilities cannot be derived explicitly. Thus, we employ an improved maximum entropy method with several well-known constraints to estimate the probability distributions of system size and the expected waiting time in the system. By a comparative analysis between the exact and approximate results, we may demonstrate that the improved maximum entropy method is accurate enough for practical purpose, and it is a useful method for solving complex queueing systems.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

In this paper, we consider an unreliable server in an M/G/1 queue operating under the $(p,N)$-policy with a second optional service (here abbreviated as SOS) and general startup times. An unreliable server means that the service is typically subject to unpredictable breakdowns. We elaborate an information theoretic technique based on the principle of maximum entropy to give an alternative solution for deriving probability distributions in this queueing model. We call that the policy is a $(p,N)$-policy if it prescribes the following conditions: (i) turn the server off when the system is empty, (ii) turn the server on if there are $N/N \geq 1$ or more customers are present, (iii) if the server is turned off and the number of customers in the system reaches $N$, turn the server on with probability $p$ and leave the server off with probability $(1 - p)$, and (iv) do not turn the server at other epochs. If the server finds at least $N$ customers present in the system, it starts to provide first essential service (here abbreviated as FES) for the waiting customers whenever he completes its startup. In other words, the $(p,N)$-policy is to control the server randomly at the arrival epoch of the $N$th customer finds that the server is idle. If the probability $p$ is one, then we have $N$-policy introduced by Yadin and Naor (1963). In case $p = 0$, we have the $(N + 1)$-policy. An M/G/1 queue involving the randomized server control problem has been treated by Feinberg and Kim (1996). They considered either $(p,N)$- or $(N,p)$-policy M/G/1 queue with a removable sever at first and performed the optimal control policy is of the randomized form. Subsequently, Kim and Moon (2006) considered the system with the $(p,N)$-policy, exploit its properties and found the optimal values of $T$ and $p$ for a constrained problem. Lately, Ke, Ko, and Sheu (2008) utilized bootstrap methods to investigate the estimation of the expected busy period of an M/G/1 queueing system under $(p,N)$-policy.

One of the most significant regions of queueing problem is the control of queue, and have studied extensive by many researchers. Yadin and Naor (1963) first introduced the concept of an $N$-policy which turns the server on whenever $N$ $(N \geq 1)$ or more customers are present, turns the server off only when the system is empty. The server startup corresponds to the preparatory work of the server before starting the service. In some actual situations, the server often needs a startup time before providing service. Exact steady-state solutions of the $N$ policy M/M/1 queue with exponential startup times were first derived by Baker (1973). Borthakur, Medhi, and Gohain (1987) extended Baker’s model to general startup times. Wang (2003) developed the exact steady-state solutions of the $N$ policy M/M/1 queue with server breakdowns and exponential startup times. The $N$-policy M/G/1 queue with startup times was

* Corresponding author.

E-mail address: khwang@math.nchu.edu.tw (K.-H. Wang).
investigated by several authors such as Medhi and Templeton (1992), Takagi (1993), Lee and Park (1997), etc. Ke (2003) analyzed the N policy M/G/1 queueing system with server vacations, startup and breakdowns. He developed the probability generating function of the queue size when the server begins performing startup and also derived important system characteristics. Recently, Wang, Wang, and Pearn (2007) focused mainly on performing a sensitivity analysis for the N-policy with server breakdowns and general startup times.

In many real service systems, one encounters numerous examples of the queueing situation where all arrivals require the main service and only some may require the subsidiary service provided by the server. Madan (2000) was the first to study an M/G/1 queue with SOS in which the first essential service time obeys a general distribution but second optional service time follows an exponential distribution. He also cited some important examples in daily life. Medhi (2002) extended Madan’s model (Madan, 2000) that the second optional service time follows a general distribution. Al-Jaraha and Madan (2003) generalized Madan’s work in the sense that they assumed that both first essential service time and second optional service time are general with different distribution functions. Based on the supplementary variable technique, Wang (2004) studied the reliability behavior in an M/G/1 queue with SOS and an unreliable server. Recently, Wang and Zhao (2007) considered a discrete-time Geo/G/1 retrial queue with an unreliable server and SOS. Some performance measures of the system in steady state and explicit formulae for the stationary distribution are developed in their work.

In a stochastic context, little is known analytically about the behaviors of queue length distributions of a randomized server control queueing system. When exact methods of solution are not known, we frequently make use of numerical solution methods. One elegant approach for this is given by the information theoretic technique, which is based on the principle of maximum entropy, to provide a self-contained method of inference for obtaining an unknown and unique probability distribution. In other words, this method is applied to estimate probability distributions, which consists of maximizing entropy function subject to the available mean constraints. El-Affendi and Kouvatmos (1983) presented the maximum entropy formalism to analyze the M/G/1 and G/M/1 queues. Based on the maximum entropy principle, Artalejo and Lopez-Herrero (2004) investigated the probability density function of busy period under some controllable M/G/1 queueing models. Wang, Wang, and Pearn (2005) used maximum entropy analyses to study the N policy M/G/1 queueing system with server breakdowns and general startup times. Recently, Ke and Lin (2006) studied the M^N/G/1 queueing system with an unreliable server and delaying vacations. They derived the approximate steady-state probability distribution of the queue length as well. To the best of our knowledge, there has been no research that investigates a randomized controllable queueing system with SOS and startup times by the maximum entropy principle. Our work is motivated by such works and employ maximum entropy method to estimate the queue length distribution for the (p,N)-policy M/G/1 queue with server breakdowns, SOS and startup times.

The purpose of this paper is fourfold. Firstly, we develop some exact and important system performance measures for the (p,N)-policy M/G/1 queue with server breakdowns, SOS and startup times. Secondly, we construct an improved maximum entropy function for this queueing system. Thirdly, the improved maximum entropy solutions are developed through the Lagrange’s method. Thirdly, we obtain the approximate expected waiting time in the system and the exact expected waiting time in the system. Finally, we perform a comparative analysis between approximate results obtained through the improved maximum entropy method and exact results obtained from the convex combination property.

2. The mathematical model

In this paper, we consider the (p,N) M/G/1 queue with the following specifications. It is assumed that customers arrive according to a Poisson process with rate \( \lambda \). Arriving customers form a single waiting line at a server based on the order of their arrivals; that is, in a first-come, first-served (FCFS) discipline. A single server is required to serve all arriving customers for the first essential service (FES), denoted by \( S_1 \). As soon as FES of a customer is completed, a customer may leave the system with probability 1 - \( \theta \) or may opt for SOS, denoted by \( S_2 \), with probability \( \theta \) (0 < \( \theta \) ≤ 1), at the completion of which the customer departs from the system and the next customer, if any, from the queue is taken up for his FES. The service times \( S_1, S_2 \) of two channels are independent and identically distributed (i.i.d.) random variables obeying a general distribution function \( S_i(t) \) (t ≥ 0), \( i = 1, 2 \), mean service time \( \mu_i \), \( i = 1, 2 \), startup times are \( \mu_{i2} \), \( i = 1, 2 \). Laplace-Stieltjes transforms (LST) \( f_{S_i}(s) \) = 1, 2, and the kth moment \( E[S_i^k] \), \( k = 1, 1 \), \( i = 1, 2 \), where the sub-index \( i = 1 \) (respectively \( i = 2 \)) denote the FES (respectively the SOS). Further, the same server is assumed to serve both service channels. Therefore, a total service time is provided to a customer is defined as:

\[
S = \begin{cases} S_1 + S_2, & \text{with probability } \theta, \\ S_1, & \text{with probability } (1 - \theta), 
\end{cases}
\]

and its LST \( f_S(s) = (1 - \theta)f_{S_1}(s) + \theta f_{S_2}(s) \) with the first moments of \( S \) are:

\[
E[S] = E[S_1] + \theta E[S_2] = \mu_1 + \theta \mu_2, \quad (1)
\]

\[
E[S^2] = E[S_1^2] + 2\theta E[S_1]E[S_2] + \theta E[S_2^2]. \quad (2)
\]

When the server is working, it may meet unpredictable breakdowns but is immediately repaired. We assume that a server's breakdown time has an exponential distribution with rate \( \gamma_1 \) in the FES channel. In the SOS channel, the server fails at an exponential rate \( \gamma_2 \). When the server fails, it is immediately repaired at a repair facility. The repair times of FES and SOS channels are independent general distributions with distribution functions \( R_1(t), R_2(t), (t \geq 0) \), mean repair times \( \mu_{r1}, \mu_{r2} \) and the kth moment \( E[R_1^k], E[R_2^k], k \geq 1 \). Although no service occurs during the repair period of the server, customers continue to arrive following a Poisson process. Once the failed server is repaired, it immediately returns to serve a customer until the system is empty.

The idle server operates the \( (p,N) \)-policy when there are \( N \) customers accumulated in the system. He requires a startup time with random length before starting FES. Again, the startup times are independent and identically distributed random variables obeying a general distribution function \( U(t) \) (t ≥ 0), mean startup time \( \mu_{l0} \) and the kth moment \( E[U^k], k \geq 1 \). As soon as the server completes startup, it begins serving the waiting customers until the system is empty. Let us suggest to the usual independence assumptions between inter-arrival times, service times, inter-breakdown times, startup times and repair times. Consequently, We will present this queueing model as the \( (p,N) \)-policy \( M/(G,G), \) \( (G,G) \), G/1 queue, where the second and third symbols denote service time distribution for FES and SOS channels, respectively. The fourth and fifth symbols denote the repair time distributions for FES and SOS channels, respectively. The sixth symbol is the startup time distribution.

3. System performance measures

Let \( H_1 \) and \( H_2 \) be a random variable representing the completion time of FES and SOS, respectively. The completion time of a customer includes both the service time of a customer and the repair time of a server. Using the known results of Wang and Ke (2002),
we get the first two moments of the completion time distribution for the first essential channel and second optional channel:

\[E[H_1] = \mu_b (1 + \alpha_1 \mu_b), \quad i = 1, 2, \]  
\[E[H_2^i] = \left(1 + \alpha_1 \mu_b\right)^2 E[S_1^i] + \alpha_1 \mu_b E[R_1^i], \quad i = 1, 2. \]  

We denote by \( I_{N}, U_{N}, B_{N}, D_{N} \), idle, startup, busy, breakdown periods for the \((p, N)\)-policy \(M|G, G, G, G, G, G, G| 1\) queue, respectively. Suppose that \( C_N \) is a busy cycle, which is a sum of idle, startup, busy, breakdown periods. Applying the results of Wang et al. (2007), we have:

\[E[I_N] = \frac{N}{\lambda}, \]  
\[E[U_N] = \frac{\rho_0}{\lambda}, \]  
\[E[B_N] = \frac{E[S](N + \rho_0)}{1 - \rho_H}, \]  
\[E[D_N] = \frac{1}{1 - \rho_H} \left(\alpha_1 \mu_b \mu_b + \theta \alpha_2 \mu_b \mu_b\right)(N + \rho_0)/1 - \rho_H. \]

We denote by \( N_1, N_2, N_3 \), idle, startup, busy, breakdown periods for the \((p, N)\)-policy \(M|G, G, G, G, G, G, G, G| 1\) queue. And let \( C_N \), be a busy cycle for the \((p, N)\)-policy \(M|G, G, G, G, G, G, G, G| 1\) queue. Based on the arguments of Feinberg and Kim (1996), it shows that the system performance measures for the \((p, N)\)-policy queue is a convex combination of the performance measures for the \(N\)-policy queue and the performance measures for the \((N + 1)\)-policy queue. Using the above formulas (5)–(9), we can obtain:

\[E[C_{p,N}] = pE[C_N] + (1 - p)E[C_{N+1}] = \frac{N + 1 - p + \rho_0}{\lambda(1 - \rho_H)}. \]

### 3.1. The long-run fraction of time measures

We will develop the maximum entropy solutions for steady-state probabilities of the \((p, N)\)-policy \(M|G, G, G, G, G, G, G, G| 1\) queue. Steady-state probabilities \(P(n), P_1(n), P_2(n), Q_1(n)\) and \(Q_2(n)\) for the entropy formalism are defined as follows:

\[P(n) \equiv \text{probability that there are } n \text{ customers in the system when the server is turned off, where } n = 0, 1, 2, \ldots, N - 1, N \]

\[P_1(n) \equiv \text{probability that there are } n \text{ customers in the system when the server is startup, where } n = N - 1, N - 1, \ldots \]

\[P_2(n) \equiv \text{probability that there are } n \text{ customers in the queue excluding the one being provided SOS, and the server is in operation, where } n = 1, 2, 3, \ldots \]

\[Q_1(n) \equiv \text{probability that there are } n \text{ customers in the queue excluding the one being provided SOS, and the server is in operation} \]

\[Q_2(n) \equiv \text{probability that there are } n \text{ customers in the queue excluding the one being provided SOS, and the server is in operation} \]

From Eqs. (13)–(17), we can easily obtain the following probabilities for the \((p, N)\)-policy \(M|G, G, G, G, G, G, G, G| 1\) queue.

The probability that the server is idle given by

\[\sum_{n=0}^{N} P_1(n) = \frac{E[I_{p,N}]}{E[C_{p,N}]} = \frac{(N + 1 - p)(1 - \rho_0)}{N + 1 - p + \rho_0}. \]

The probability that the server is startup given by

\[\sum_{n=0}^{N} P_2(n) = \frac{E[U_{p,N}]}{E[C_{p,N}]} = \frac{\rho_0(1 - \rho_0)}{N + 1 - p + \rho_0}. \]

The probability that the server is busy given by

\[\sum_{n=0}^{N} P_1(n) + \sum_{n=1}^{\infty} P_2(n) = \frac{E[B_{p,N}]}{E[C_{p,N}]} = \lambda E[S] = \lambda \mu_b + \theta \lambda \mu_b. \]

The probability that the server is breakdown given by

\[\sum_{n=0}^{N} Q_1(n) + \sum_{n=1}^{\infty} Q_2(n) = \frac{E[D_{p,N}]}{E[C_{p,N}]} = \lambda \alpha_1 \mu_b \mu_b + \theta \lambda \alpha_2 \mu_b \mu_b. \]

For a start, we note that the long-run fraction of time the server is busy is FES or SOS is provided, and can be represented as \(\lambda \mu_b \) and \(\theta \lambda \mu_b \), respectively. Next, it is noticed that the long-run fraction of time the server is broken down when the FES or SOS provided, which can also be represented as \(\lambda \alpha_1 \mu_b \mu_b \) and \(\theta \lambda \alpha_2 \mu_b \mu_b \), respectively.

### 3.2. The expected number of customers in the system

Let \( T_{N,1} \) and \( T_{p,n} \) denote the cumulative amount of time that all customers spent in the system during a busy cycle for the \((N, \ldots, N)\)- and \((p, N)\)-policies \(M|G, G, G, G, G, G, G, G| 1\) queue. Following the results of Feinberg and Kim (1996), we can obtain:
From Eq. (18) and the results of Wang et al. (2007), we have:

\[ P_1(N) = \frac{(1 - p)(1 - \rho_H)}{N + 1 - p + \rho_U}. \]  

(27)

4. Improved maximum entropy results

Exact probability distributions of the \( \langle p,N \rangle \)-policy \( M/(G,G), (G,G), G/1 \) queue have not been found. Therefore, we employ the improved maximum entropy principle to estimate probability distributions of the number of customers given several known results. In this section, we will develop the improved maximum entropy solutions for the steady-state probabilities of the \( \langle p,N \rangle \)-policy \( M/(G,G), (G,G), G/1 \) queue.

4.1. The improved maximum entropy model

In order to derive the approximate steady-state probabilities \( P_i(n), P_i(n) \) \( (i = 1,2), Q_i(n) \) \( (i = 1,2) \), we formulate the maximum entropy model in the following. Because that the exact results for \( P_i(0) \) and \( P_i(N) \) are known, the improved entropy function \( Y \) of the \( \langle p,N \rangle \)-policy \( M/(G,G), (G,G), G/1 \) queue can be formed as:

\[
Y = -\sum_{n=0}^{\infty} P_i(n) \ln P_i(n) - \sum_{n=1}^{\infty} P_i(n) \ln P_i(n) \\
- \sum_{n=1}^{\infty} P_i(n) \ln P_i(n) - \sum_{n=1}^{\infty} Q_i(n) \ln Q_i(n) \\
- \sum_{n=1}^{\infty} Q_i(n) \ln Q_i(n). 
\]  

(28)

The improved maximum entropy solutions for the \( \langle p,N \rangle \)-policy \( M/(G,G), (G,G), G/1 \) queue are obtained by maximizing Eq. (28) subject to the following six constraints, written as:

1. The probability that the server is startup:

\[
\sum_{n=0}^{\infty} P_i(n) = \frac{\rho_0(1 - \rho_H)}{N + 1 - p + \rho_U} = \Pi \rho_0(1 - \rho_H),
\]  

(29)

where \( \Pi = 1/(N + 1 - p + \rho_U) \).

2. The probability that the server is busy of providing FES:

\[
\sum_{n=1}^{\infty} P_i(n) = \lambda \mu_{q_1} = \rho_1.
\]  

(30)

3. The probability that the server is busy of providing SOS:

\[
\sum_{n=1}^{\infty} P_i(n) = \lambda \mu_{q_2} = \rho_2.
\]  

(31)

4. The probability that the server is broken down when FES is provided:

\[
\sum_{n=1}^{\infty} Q_i(n) = \rho_2 \sigma_1 \mu_{b_1}.
\]  

(32)

5. The probability that the server is broken down when SOS is provided:

\[
\sum_{n=1}^{\infty} Q_i(n) = \rho_2 \sigma_2 \mu_{b_2}.
\]  

(33)

6. The expected number of customers in the system when the server is not idle:

\[
\sum_{n=0}^{\infty} nP_i(n) + \sum_{n=1}^{\infty} nP_i(n) + \sum_{n=1}^{\infty} nP_i(n) + \sum_{n=1}^{\infty} nQ_i(n)
\]

\[
+ \sum_{n=1}^{\infty} nQ_i(n) = L_{p,N} - L_1,
\]  

(34)

where \( L_1 \) is the expected length of customers as the server is idle and can be expressed as follows:

\[
L_1 = \frac{N(N - 1)}{2} P_i(0) + \frac{N(N + 1 - 2p)(1 - \rho_H)}{2(N + 1 - p + \rho_U)}.
\]  

(35)
In Eqs. (29)–(34), Eq. (29) is multiplied by $\delta_1$, Eq. (30) is multiplied by $\delta_2$, Eq. (31) is multiplied by $\delta_3$, Eq. (32) is multiplied by $\delta_4$, Eq. (33) is multiplied by $\delta_5$, Eq. (34) is multiplied by $\delta_6$. Thus, the Lagrangian function $y$ is given by:

$$
y = - \sum_{n=1}^{\infty} P_n(n) \ln P_n(n) - \sum_{n=1}^{\infty} P_i(n) \ln P_i(n)
- \sum_{n=1}^{\infty} Q_n(n) \ln Q_n(n) - \delta_1 \left[ \sum_{n=1}^{\infty} P_n(n) - \Pi \rho_1(1 - \rho_n) \right]
- \delta_2 \left[ \sum_{n=1}^{\infty} P_i(n) - \rho_1 \right]
- \delta_3 \left[ \sum_{n=1}^{\infty} P_2(n) - \theta_2 \right]
- \delta_4 \left[ \sum_{n=1}^{\infty} Q_1(n) - \rho_1 x \mu_{R_1} \right]
- \delta_5 \left[ \sum_{n=1}^{\infty} Q_2(n) - \theta_2 x \mu_{R_2} \right]
- \delta_6 \left[ \sum_{n=1}^{\infty} nP_n(n) + \sum_{n=1}^{\infty} nP_i(n) + \sum_{n=1}^{\infty} nP_2(n) + \sum_{n=1}^{\infty} nQ_1(n)
+ \sum_{n=1}^{\infty} nQ_2(n) - L_p N + L_i \right],
$$

where $\delta_1$ to $\delta_6$ are the Lagrange multipliers corresponding to constraints (29)–(34), respectively.

### 4.2. The improved maximum entropy solutions

To find the improved maximum entropy solutions $P_3(n), P_i(n)$ ($i = 1, 2$) and $Q_1(n)$ ($i = 1, 2$), maximizing in (28) subject to constraints (29)–(34) is equivalent to maximizing (36). The improved maximum entropy solutions are obtained by taking the partial derivatives of $y$ with respect to $P_n(n), P_i(n)(i = 1, 2), Q_1(n)(i = 1, 2)$ and setting the results equal to zero, namely:

$$
\frac{\partial y}{\partial P_3(n)} = - \ln P_3(n) - 1 - \delta_1 - \delta_6 n = 0, \quad n = N, N + 1, \ldots
$$

$$
\frac{\partial y}{\partial P_1(n)} = - \ln P_1(n) - 1 - \delta_2 - \delta_6 n = 0, \quad n = 1, 2, \ldots
$$

$$
\frac{\partial y}{\partial P_2(n)} = - \ln P_2(n) - 1 - \delta_3 - \delta_6 n = 0, \quad n = 1, 2, \ldots
$$

$$
\frac{\partial y}{\partial Q_1(n)} = - \ln Q_1(n) - 1 - \delta_4 - \delta_6 n = 0, \quad n = 1, 2, \ldots
$$

$$
\frac{\partial y}{\partial Q_2(n)} = - \ln Q_2(n) - 1 - \delta_5 - \delta_6 n = 0, \quad n = 1, 2, \ldots
$$

It follows from Eqs. (37)–(41) that:

$$
P_3(n) = e^{-(1+\delta_1)e^{-\delta_6 n}}, \quad n = N, N + 1, \ldots
$$

$$
P_1(n) = e^{-(1+\delta_2)e^{-\delta_6 n}}, \quad n = 1, 2, \ldots
$$

$$
P_2(n) = e^{-(1+\delta_3)e^{-\delta_6 n}}, \quad n = 1, 2, \ldots
$$

$$
Q_1(n) = e^{-(1+\delta_4)e^{-\delta_6 n}}, \quad n = 1, 2, \ldots
$$

$$
Q_2(n) = e^{-(1+\delta_5)e^{-\delta_6 n}}, \quad n = 1, 2, \ldots
$$

Let $\omega_i = e^{-(1+\delta_i)}$ for $1 \leq i \leq 5$, and $\omega_6 = e^{-\delta_6}$. We transform Eqs. (42)–(46) in terms of $\omega_i (1 \leq i \leq 6)$ given by:

$$
P_3(n) = \omega_4 \omega_6^n, \quad n = N, N + 1, \ldots
$$

$$
P_1(n) = \omega_2 \omega_6^n, \quad n = 1, 2, \ldots
$$

$$
P_2(n) = \omega_3 \omega_6^n, \quad n = 1, 2, \ldots
$$

$$
Q_1(n) = \omega_4 \omega_6^n, \quad n = 1, 2, \ldots
$$

$$
Q_2(n) = \omega_2 \omega_6^n, \quad n = 1, 2, \ldots
$$

Substituting Eqs. (47)–(51) into Eqs. (29)–(33), respectively, yields:

$$
o_1 = \frac{\Pi \rho_1(1 - \rho_2)(1 - \omega_6)}{\omega_6},
$$

$$
o_2 = \frac{\rho_1(1 - \omega_6)}{\omega_6},
$$

$$
o_3 = \frac{\theta_2(1 - \omega_6)}{\omega_6},
$$

$$
o_4 = \frac{\rho_1 \lambda X \mu_{R_1}(1 - \omega_6)}{\omega_6},
$$

$$
o_5 = \frac{\theta_2 \lambda X \mu_{R_2}(1 - \omega_6)}{\omega_6}.
$$

Substituting Eqs. (47)–(51) into Eq. (34) and taking the algebraic manipulations, we obtain:

$$
o_6 = \frac{\Pi \rho_1(1 - \rho_2) + \rho_6}{L_p N - \Pi(1 - \rho_n) \left[ \frac{N(N+1-2\omega_6)}{2} + (N - 1)\rho_1 \right]}.
$$

Substituting Eqs. (52)–(57) into Eqs. (47)–(51), respectively, we finally get:

$$
P_3(n) = \frac{\Pi \rho_1(1 - \rho_2)(1 - \omega_6)\omega_6^n}{N(N+1-2\omega_6)}, \quad n = N, N + 1, \ldots
$$

$$
P_1(n) = \frac{\rho_1(1 - \omega_6)\omega_6^n}{N + 1, \ldots}
$$

$$
P_2(n) = \frac{\theta_2(1 - \omega_6)\omega_6^n}{N + 1, \ldots}
$$

$$
Q_1(n) = \frac{\rho_1 \lambda X \mu_{R_1}(1 - \omega_6)\omega_6^n}{N + 1, \ldots}
$$

$$
Q_2(n) = \frac{\theta_2 \lambda X \mu_{R_2}(1 - \omega_6)\omega_6^n}{N + 1, \ldots}
$$

### 5. The exact and approximate expected waiting time in the system

In this section, we first derive the exact expected waiting time in the system by using Little’s formula. Through the maximum entropy principle, the approximate formulae of the expected waiting time in the system for the $(p,N)$-policy M/$(G,G),(G,G)$, GI/1 queue is developed.

#### 5.1. The exact expected waiting time in the system

Let $W_s(N), W_s(N+1)$ and $W_s(p,N)$ denote the exact expected waiting time in the system for the $N$, $(N+1)$- and $(p,N)$-policies, respectively. Using Eqs. (10) and (24) in Little’s formula, we see that:

$$
W_s(N) = \frac{L_p}{\lambda} = \frac{1}{N + \rho_2} \left[ \frac{N(N+1)}{2 \lambda} + N \mu + \frac{\lambda E[U^2]}{2} \right] + \frac{L_2}{\lambda}.
$$

$$
W_s(N+1) = \frac{L_p}{\lambda} = \frac{1}{N + 1 + \rho_2} \left[ \frac{N(N+1)}{2 \lambda} + (N+1) \mu + \frac{\lambda E[U^2]}{2} \right] + \frac{L_2}{\lambda}.
$$

$$
W_s(p,N) = \frac{L_p}{\lambda} = \frac{1}{N + 1 - p + \rho_2}
\times \left[ \frac{N(N+1-2p)}{2 \lambda} + (N + 1 - p) \mu + \frac{\lambda E[U^2]}{2} \right] + \frac{L_2}{\lambda}.
$$

From Feinberg and Kim (1996), we know that $W_s(p,N)$ is a convex combination of $W_s(N)$ and $W_s(N+1)$. It follows that:

$$
W_s(p,N) = \frac{p(N+\rho_2)}{N+1-p+\rho_2} W_s(N)
+ \frac{1-p(N+\rho_2)}{N+1-p+\rho_2} W_s(N+1).
$$
Substituting Eqs. (63) and (64) into Eq. (66), we have the same result shown in Eq. (65). Thus, we demonstrate that the relationships given by Eqs. (65) and (66) are seen to hold.

5.2. The approximate expected waiting time in the system

The idle state, the startup state, the busy state, and the repair state are defined as follows:

1. Idle state 1 denoted by \( I_1 \): the server is turned off, and the number of customers waiting in the system is less than or equal to \( N - 1 \).
2. Idle state 2 denoted by \( I_2 \): the server is turned off, and the number of customers waiting in the system is equal to \( N \).
3. Startup state denoted by \( U \): the server begins startup, and the number of customers waiting in the system is greater than or equal to \( N \).
4. Busy state when \( FES \) is provided denoted by \( B_1 \): the server is busy and provides \( FES \) to a customer.
5. Busy state when \( SOS \) is provided denoted by \( B_2 \): the server is busy and provides \( SOS \) to a customer.
6. Repair state when \( FES \) is provided denoted by \( R_1 \): the server is broken down when \( FES \) is provided and being repaired.
7. Repair state when \( SOS \) is provided denoted by \( R_2 \): the repair state is broken down when \( SOS \) is provided and being repaired.

We wish to find the expected waiting time of an arbitrary customer \( C \) at the state \( I_1, I_2, U, B_1, B_2, R_1 \) and \( R_2 \). Suppose an arbitrary customer \( C \) finds \( n \) customers waiting in the queue for service in front of him, while the system is at any one of the states \( I_1, I_2, U, B_1, B_2, R_1 \) and \( R_2 \) are described, respectively, as follows:

1. In idle state \( I_1 \): Note that the idle state immediately is switched to startup state after an arbitrary customer \( C \) arrives and \( n \) customers in front of him are waiting for service. The server will begin startup after \( (N - n - 1) \) customers arrive with probability \( p \) or after \( (N - n) \) customers arrive with probability \( (1 - p) \) in the system. Thus customer \( C \) will be served until \( (N - n - 1) \) customers arrive with probability \( p \) or \( (N - n) \) customers arrive with probability \( (1 - p) \), and \( n \) customers in front of him are waiting for service. Hence, customer \( C \) must wait (i) the mean residual idle time, (ii) the service time of \( n \) customers in the system and (iii) the startup time before providing \( FES \). From the inferences of (i)–(iii), the expected waiting time of customer \( C \) at the idle state \( I_1 \) is

\[
\frac{(N - n - 1)p}{\lambda} + \frac{(N - n)(1 - p)}{\lambda} + \mu_U + nE[S] = \frac{N - n - p}{\lambda} + \mu_U + nE[S].
\]

2. In idle state \( I_2 \): The server will begin startup when there are \( N \) customers present in the system. Thus customer \( C \) will be served when no customers in front of him for waiting for service. The expected waiting time of customer \( C \) at the idle state \( I_2 \) is \( \mu_0 + nE[S] \).

3. In startup state \( U \): We derive the expected waiting time of customer \( C \) at the startup state in the following. Let us define:

\[
F_U(t) = \Pr(U_1(t) \leq t) = \frac{1}{\mu_U} \int_0^t [1 - D(x)]dx,
\]

where \( D(x) \) is the c.d.f. of startup time. Let \( E[U_1] \) be the mean remaining startup time. It implies that \( E[U_1] = E[U^2]/2\mu_U \). Thus we obtain the expected waiting time of customer \( C \) at the startup state is

\[
nE[S] + E[U^2]/2\mu_U.
\]

4. In busy states \( B_1 \) and \( B_2 \): Since the server is busy and keeps working, the customer \( C \) only waits \( n \) customers who demand the server in front of him. The expected waiting time at the busy states \( B_1 \) and \( B_2 \) are \( nE[S] \).

5. In repair states \( R_1 \) and \( R_2 \): According to the same argument as (3), we have the expected waiting time of an arbitrary customer \( C \) at the repair states \( R_1 \) and \( R_2 \) are

\[
nE[S] + E[U^2]/2\mu_U \text{, and } nE[S] + E[U^2]/2\mu_U \text{, respectively.}
\]

Utilizing the listed above results, we obtain the approximate expected waiting time in the queue, \( W^*_C(p, N) \), given by

\[
W^*_C(p, N) = \sum_{n=0}^{N-1} \left( \frac{N - n - p}{\lambda} + \mu_U + nE[S] \right) P_1(0) + (\mu_U + nE[S]) P_1(N) + \sum_{n=N}^{\infty} \left( nE[S] + E[U^2]/2\mu_U \right) P_2(n) + \sum_{n=1}^{\infty} \left( nE[S] + E[U^2]/2\mu_U \right) Q_1(n) + \sum_{n=1}^{\infty} \left( nE[S] + E[U^2]/2\mu_U \right) Q_2(n).
\]

Substituting Eqs. (26), (27), (28)–(32) into Expression (66), the approximate expected waiting in the queue is given by

\[
W_p^*(p, N) = \frac{N(N + 1 - 2p)(1 - \rho_H)}{2\lambda(N + 1 - p + \rho_H)} + \frac{2\mu_U(N + 1 - p + \lambda E[U^2])}{2(N + 1 - p + \rho_H)} + \frac{p W_{\bar{S}}(p, N) + E[R_1^2] \rho_1 \rho_2 \chi_1 + \theta E[R_2^2] \rho_2 \chi_2}{2}.
\]

where the derivation of Eq. (68) is shown in Appendix. Consequently, we again use Little’s formula to obtain the approximate expected waiting time in the system as follows:

\[
W_S(p, N) = \frac{N(N + 1 - 2p)(1 - \rho_H)}{2\lambda(N + 1 - p + \rho_H)} + \frac{2\mu_U(N + 1 - p + \lambda E[U^2])}{2(N + 1 - p + \rho_H)} + \frac{p W_{\bar{S}}(p, N) + E[R_1^2] \rho_1 \rho_2 \chi_1 + \theta E[R_2^2] \rho_2 \chi_2}{2} + E[H].
\]

6. Comparative analysis between exact and approximate results

This section aims to examine the accuracy of the approximate results based on the improved maximum entropy principle. We provide numerical comparisons between the exact results and the approximate results, including various service time, startup time and repair time distribution functions. There are three subsections in the following:

1. Comparative analysis for the \( (p, N) \)-policy \( M/(M, E_2), (M, D), M \| 1 \) queue.
Comparison of exact

The relative error percentage for the $(p,N)$-policy $M/(M,D)$, $(E_2,E_3)$, $D$ queue.

Here, $M$ is an exponential distribution, $D$ is a deterministic distribution and $E_k$ is a $k$-stage Erlang distribution.

6.1. Comparative analysis for the $(p,N)$-policy $M/(M,D)$, $(E_2)$, $D$ queue

We perform a comparative analysis between the exact $W(p,N)$ and the approximate $W'_p(p,N)$ for the $(p,N)$-policy $M/(M,D)$, $(E_2)$, $D$ queue. For this queueing system, we have:

Table 1
The relative error percentage for the $(p,N)$-policy $M/(E_2)$, $(M,D)$, $M/1$ queue ($\lambda = 0.5, \mu_1 = 1.0, \mu_2 = 2.0, \gamma = 3.0$, $x_1 = 0.05, x_2 = 0.10, \beta_1 = 3.0, \beta_2 = 4.0, \theta = 0.4$).

<table>
<thead>
<tr>
<th>$N$</th>
<th>$p$</th>
<th>$p = 0.01$</th>
<th>$p = 0.05$</th>
<th>$p = 0.09$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.594</td>
<td>0.604</td>
<td>0.656</td>
<td>0.720</td>
</tr>
<tr>
<td>4</td>
<td>0.122</td>
<td>0.134</td>
<td>0.165</td>
<td>0.200</td>
</tr>
<tr>
<td>6</td>
<td>0.141</td>
<td>0.133</td>
<td>0.114</td>
<td>0.092</td>
</tr>
<tr>
<td>8</td>
<td>0.311</td>
<td>0.305</td>
<td>0.292</td>
<td>0.277</td>
</tr>
<tr>
<td>10</td>
<td>0.428</td>
<td>0.424</td>
<td>0.414</td>
<td>0.404</td>
</tr>
<tr>
<td>12</td>
<td>0.515</td>
<td>0.512</td>
<td>0.504</td>
<td>0.488</td>
</tr>
<tr>
<td>14</td>
<td>0.581</td>
<td>0.579</td>
<td>0.573</td>
<td>0.567</td>
</tr>
<tr>
<td>16</td>
<td>0.634</td>
<td>0.632</td>
<td>0.627</td>
<td>0.622</td>
</tr>
<tr>
<td>18</td>
<td>0.676</td>
<td>0.675</td>
<td>0.671</td>
<td>0.667</td>
</tr>
<tr>
<td>20</td>
<td>0.712</td>
<td>0.710</td>
<td>0.707</td>
<td>0.704</td>
</tr>
</tbody>
</table>

Table 2
Comparison of exact $W(p,N)$ and approximate $W'_p$ for the $(p,N)$-policy $M/(E_2)$, $(M,D)$, $M/1$ queue ($N = 8$).

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$W(p,N)$</th>
<th>$W'_p(p,N)$</th>
<th>RE%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.1</td>
<td>0.01</td>
<td>0.1</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.01</td>
<td>0.1</td>
</tr>
<tr>
<td>0.4</td>
<td>0.01</td>
<td>0.01</td>
<td>0.1</td>
</tr>
<tr>
<td>0.6</td>
<td>0.01</td>
<td>0.01</td>
<td>0.1</td>
</tr>
<tr>
<td>0.8</td>
<td>0.01</td>
<td>0.01</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 3
The relative error percentage for the $(p,N)$-policy $M/(M,D)$, $(E_1,E_3)$, $D$ queue ($\lambda = 0.5, \mu_1 = 1.0, \mu_2 = 2.0, \gamma = 3.0$, $x_1 = 0.05, x_2 = 0.10, \beta_1 = 3.0, \beta_2 = 4.0, \theta = 0.4$).

<table>
<thead>
<tr>
<th>$N$</th>
<th>$p$</th>
<th>$p = 0.01$</th>
<th>$p = 0.05$</th>
<th>$p = 0.09$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.855</td>
<td>0.879</td>
<td>0.941</td>
<td>1.016</td>
</tr>
<tr>
<td>4</td>
<td>0.314</td>
<td>0.329</td>
<td>0.364</td>
<td>0.405</td>
</tr>
<tr>
<td>6</td>
<td>0.007</td>
<td>0.017</td>
<td>0.040</td>
<td>0.065</td>
</tr>
<tr>
<td>8</td>
<td>0.189</td>
<td>0.183</td>
<td>0.167</td>
<td>0.150</td>
</tr>
<tr>
<td>10</td>
<td>0.326</td>
<td>0.321</td>
<td>0.310</td>
<td>0.298</td>
</tr>
<tr>
<td>12</td>
<td>0.426</td>
<td>0.420</td>
<td>0.414</td>
<td>0.405</td>
</tr>
<tr>
<td>14</td>
<td>0.503</td>
<td>0.500</td>
<td>0.494</td>
<td>0.487</td>
</tr>
<tr>
<td>16</td>
<td>0.564</td>
<td>0.562</td>
<td>0.556</td>
<td>0.551</td>
</tr>
<tr>
<td>18</td>
<td>0.613</td>
<td>0.611</td>
<td>0.607</td>
<td>0.603</td>
</tr>
<tr>
<td>20</td>
<td>0.654</td>
<td>0.653</td>
<td>0.649</td>
<td>0.645</td>
</tr>
</tbody>
</table>
Firstly, we fix accuracy of the approximate values is assessed by the relative error:

\[
RE = \left| W(p, N) - W(p, N) \right| \times 100\%.
\]

The relative error percentage for the \((p, N)\)-policy \(M(M_i, E, \gamma, (M, D))/M/1\) queue under various values \(p\) and \(N\) are shown in Table 1. We observe from Table 1 that (i) for fixed \(p\), the relative error percentage decreases when \(N\) ranges from 2 to 6 and increases when \(N\) ranges from 8 to 20; (ii) if \(N\) is from 2 to 4 and fixed it, the relative error percentage increases in \(p\); (iii) if \(N\) is from 6 to 20 and fixed it, the relative error percentage decreases in \(p\); (iv) the relative error percentage in Table 1 is below 1%.

Next, we set \(N = 8\) and consider the different values \(p = 0.2, 0.5\) and 0.8. Choosing the various values of \(\lambda, (\mu_1, \mu_2), (\chi_1, \chi_2), (\beta_1, \beta_2), \gamma\) and \(\theta\). The numerical results are obtained by considering the following six cases:

### Table 4
Comparison of exact \(W(p, N)\) and approximate \(W(p, N)\) for the \((p, N)\)-policy \(M(M, D, (E, E))/M/1\) queue \((N = 8)\).

<table>
<thead>
<tr>
<th>(p)</th>
<th>(W(p, N))</th>
<th>(W(p, N))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\lambda)</td>
<td>(\mu_1), (\mu_2), (\chi_1), (\chi_2), (\beta_1), (\beta_2), (\gamma), (\theta)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Case 1: (\lambda = 0.5, (\mu_1, \mu_2) = (1.0, 2.0), (\chi_1, \chi_2) = (0.05, 0.01), (\beta_1, \beta_2) = (3.0, 4.0), \gamma = 3.0, \theta = 0.4)</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>0.206</td>
<td>0.207</td>
</tr>
<tr>
<td>0.5</td>
<td>0.196</td>
<td>0.197</td>
</tr>
<tr>
<td>0.8</td>
<td>0.186</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>Case 2: (\lambda = 0.5, (\mu_1, \mu_2) = (1.0, 2.0), (\chi_1, \chi_2) = (0.05, 0.01), (\beta_1, \beta_2) = (3.0, 4.0), \gamma = 3.0, \theta = 0.4)</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>0.194</td>
<td>0.195</td>
</tr>
<tr>
<td>0.5</td>
<td>0.184</td>
<td>0.185</td>
</tr>
<tr>
<td>0.8</td>
<td>0.175</td>
<td>0.176</td>
</tr>
<tr>
<td></td>
<td>Case 3: (\lambda = 0.5, (\mu_1, \mu_2) = (1.0, 2.0), (\chi_1, \chi_2) = (0.05, 0.01), (\beta_1, \beta_2) = (3.0, 4.0), \gamma = 3.0, \theta = 0.4)</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>0.192</td>
<td>0.193</td>
</tr>
<tr>
<td>0.1</td>
<td>0.186</td>
<td>0.187</td>
</tr>
<tr>
<td>0.15</td>
<td>0.180</td>
<td>0.181</td>
</tr>
<tr>
<td></td>
<td>Case 4: (\lambda = 0.5, (\mu_1, \mu_2) = (1.0, 2.0), (\chi_1, \chi_2) = (0.05, 0.01), (\beta_1, \beta_2) = (3.0, 4.0), \gamma = 3.0, \theta = 0.4)</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.190</td>
<td>0.191</td>
</tr>
<tr>
<td>0.3</td>
<td>0.187</td>
<td>0.189</td>
</tr>
<tr>
<td>0.4</td>
<td>0.186</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>Case 5: (\lambda = 0.5, (\mu_1, \mu_2) = (1.0, 2.0), (\chi_1, \chi_2) = (0.05, 0.01), (\beta_1, \beta_2) = (3.0, 4.0), \gamma = 3.0, \theta = 0.4)</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>0.191</td>
<td>0.192</td>
</tr>
<tr>
<td>0.2</td>
<td>0.188</td>
<td>0.189</td>
</tr>
<tr>
<td>0.3</td>
<td>0.187</td>
<td>0.188</td>
</tr>
<tr>
<td>0.4</td>
<td>0.187</td>
<td>0.188</td>
</tr>
<tr>
<td>0.5</td>
<td>0.188</td>
<td>0.188</td>
</tr>
<tr>
<td>0.6</td>
<td>0.189</td>
<td>0.189</td>
</tr>
<tr>
<td>0.7</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td>0.8</td>
<td>0.191</td>
<td>0.192</td>
</tr>
<tr>
<td>0.9</td>
<td>0.192</td>
<td>0.193</td>
</tr>
<tr>
<td>1.0</td>
<td>0.193</td>
<td>0.194</td>
</tr>
</tbody>
</table>

### Table 5
The relative error percentage for the \((p, N)\)-policy \(M(E,D), (D,E))/E_1\) queue \((\lambda = 0.5, \mu_1 = 1.0, \mu_2 = 2.0, \gamma = 3.0, \alpha_1 = 2.0, \alpha_2 = 0.05, \beta_1 = 3.0, \beta_2 = 4.0, \theta = 0.4)\).
6.2. Comparative analysis for the \((p, N)\)-policy \(M/(M, D), (E_2, E_3), D/1\) queue

We perform a comparative analysis between the exact \(W_p(p, N)\) and the approximate \(W_i(p, N)\) for the \((p, N)\)-policy \(M/(M, D), (E_2, E_3), D/1\) queue. For this queueing system, we have:

\[
W_p(p, N) \quad | \quad W_i(p, N) \quad | \quad \text{RE(\%)}
\]

<table>
<thead>
<tr>
<th>(\lambda )</th>
<th>( (\mu_1, \mu_2) )</th>
<th>( (x_1, x_2) )</th>
<th>( \gamma )</th>
<th>( \theta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>( (1.0, 2.0) )</td>
<td>( (0.05, 0.10) )</td>
<td>( (3.0, 4.0) )</td>
<td>0.4</td>
</tr>
<tr>
<td>0.5</td>
<td>( (1.0, 3.0) )</td>
<td>( (0.05, 0.10) )</td>
<td>( (3.0, 4.0) )</td>
<td>0.4</td>
</tr>
<tr>
<td>0.8</td>
<td>( (1.0, 3.0) )</td>
<td>( (0.05, 0.10) )</td>
<td>( (3.0, 4.0) )</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Table 6

Comparison of exact \(W_p(p, N)\) and approximate \(W_i(p, N)\) for the \((p, N)\)-policy \(M/(M, D), (E_2, E_3), D/1\) queue (\(N = 8\)).

The relative error percentage for the \((p, N)\)-policy \(M/(M, D), (E_2, E_3), D/1\) queue under various values of \(p\) and \(N\) is shown in Table 3. It reveals that (i) for fix \(p\), the relative error percentage decreases when \(N\) ranges from 2 to 6 and increases when \(N\) ranges from 8 to 20; (ii) if \(N\) is from 2 to 6 and fixed it, the relative error percentage increases in \(p\); (iii) if \(N\) is from 8 to 20 and fixed it, the relative error percentage decreases in \(p\); (iv) the relative error percentage in Table 3 is below 1.4%.

Numerical results of the \((p, N)\)-policy \(M/(M, D), (E_2, E_3), D/1\) queue summarized in Table 4 for the above six cases. Table 4 indicates that the relative error percentages are very small (0–2.7%).

6.3. Comparative analysis for the \((p, N)\)-policy \(M/(E_3, M), (D, E_4), E_3/1\) queue

We perform a comparative analysis between the exact \(W_p(p, N)\) and the approximate \(W_i(p, N)\) for the \((p, N)\)-policy \(M/(E_3, M), (D, E_4), E_3/1\) queue. For this queueing system, we have:

\[
E[S_1] = \frac{1}{\mu_1}, \quad E[S_2] = \frac{2}{\mu_2}, \quad E[R_1] = \frac{1}{\rho_1}, \quad E[R_2] = \frac{3}{2\rho_1}, \quad E[U_1] = \frac{1}{\gamma}, \quad E[U_2] = \frac{1}{2\gamma}
\]
The relative error percentage for the \( p,N \)-policy \( M(E_2,M), \) 
\((D,E_3),E_3/1 \) queue can easily understand under various values \( p \) and \( N \) are shown in Table 5. One can easily see that (i) for fix \( p \), the relative error percentage decreases when \( N \) ranges from 2 to 16 and increases when \( N \) ranges from 18 to 20; (ii) if \( N \) is from 2 to 16 and fixed it, the relative error percentage increases in \( p \); (iii) if \( N \) is from 18 to 20 and fixed it, the relative error percentage decreases in \( p \); (iv) the relative error percentage in Table 5 is below 4.2.

Numerical results of the \( p,N \)-policy \( M(E_2,M), \) 
\((D,E_3),E_3/1 \) queue summarized in Table 6 for the above six cases. Again, it shows that the relative error percentages are very small (0–1.8%).

7. Conclusion

In this paper, we developed some important system performance measures for the \( p,N \)-policy \( M(G,G), \) \((G,G), \) \( G/1 \) queue. An elegant approach, the maximum entropy principle, is used to derive the approximate formulae for the steady-state probability distributions of the queue length. Our numerical investigations show that it is feasible to use the probability of various servers states and the expected number of customers in the system when the server is not idle. The numerical results also indicate that the relative error percentages are very small (below 4.2%). As expected, it is sufficiently accuracy to obtain the approximate estimations. Finally, based on the improved maximum entropy principle, we demonstrate that the \( p,N \)-policy \( M(G,G), \) \((G,G), \) \( G/1 \) queue is really robust to the variations of service time distribution, repair time distribution and startup time distribution functions. Consequently, this improved maximum entropy method is a useful analytic tool for approximating the solution of complex queueing systems.

Appendix

\[
W'_q(p,N) = \sum_{n=0}^{N-1} \left( \frac{N-n-p}{\lambda} + \mu_0 + nE[S] \right) \rho_1(n) + (\mu_0 + NE[S])\rho_1(n)
\]

\[
+ \sum_{n=1}^{\infty} nE[S] + \frac{E[U]}{2\mu_0} P_1(n) + \sum_{n=1}^{\infty} nE[S]P_1(n)
\]

\[
+ \sum_{n=1}^{\infty} nE[S]P_1(n) + \sum_{n=1}^{\infty} nE[S] + \frac{E[R_1^2]}{2\mu_0} Q_1(n)
\]

\[
= \sum_{n=0}^{\infty} \left( \frac{N-n-p}{\lambda} + \mu_0 + nE[S] \right) \frac{1-\rho_H}{N+1-p+\rho_U}
\]

\[
+ (\mu_0 + NE[S]) \frac{(1-p)(1-\rho_H)}{N+1-p+\rho_U} + \sum_{n=N}^{\infty} (nE[S]P_1(n)
\]

\[
+ \frac{E[U]}{2\mu_0} P_1(n) + E[S] \sum_{n=1}^{\infty} (nP_1(n) + nP_2(n) + nQ_1(n)
\]

\[
+ nQ_2(n)) + \frac{E[R_1^2]}{2\mu_0} \sum_{n=1}^{\infty} Q_1(n) + \frac{E[R_1^2]}{2\mu_0} \sum_{n=1}^{\infty} Q_2(n)
\]

References


