Ann Oper Res (2006) 141: 233–270
DOI 10.1007/s10479-006-5301-0

A *BMAP/G/1* Retrial Queue with a Server Subject to Breakdowns and Repairs

Quan-Lin Li · Yu Ying · Yiqiang Q. Zhao

Abstract In this paper, we consider a *BMAP/G/1* retrial queue with a server subject to breakdowns and repairs, where the life time of the server is exponential and the repair time is general. We use the supplementary variable method, which combines with the matrix-analytic method and the censoring technique, to study the system. We apply the *RG*-factorization of a level-dependent continuous-time Markov chain of *M/G/1* type to provide the stationary performance measures of the system, for example, the stationary availability, failure frequency and queue length. Furthermore, we use the *RG*-factorization of a level-dependent Markov renewal process of *M/G/1* type to express the Laplace transform of the distribution of a first passage time such as the reliability function and the busy period.

Keywords Retrial queues · Batch Markov arrival processes (BMAP) · Markov chains of *M/G/1* type · Markov renewal processes of *M/G/1* type · Supplementary variable method · Matrix-analytic method · Censoring technique · *RG*-factorization · Reliability

1. Introduction

A retrial system consists of a primary service facility and an orbit. Customers arrive at the service facility either from outside the system or from the orbit. Upon the arrival of a customer, if the server is busy or under repair, the arrival will join the retrial group in the orbit and try its luck again at some time later. Queueing systems with retrial customers are good mathematical models for telephone switch systems, digital cellular mobile networks and computer networks etc. During the last two decades considerable attention has been paid

Q.-L. Li
Department of Industrial Engineering, Tsinghua University, Beijing 100084, P.R. China; School of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada K1S 5B6

Y. Ying
School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, 47907, U.S.A.

Y. Q. Zhao
School of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada K1S 5B6

to studying retrial queues, which has been well documented, for example, by survey papers of Yang and Templeton (1987), Falin (1990) and Kulkarni and Liang (1997), by a book of Falin and Templeton (1997) and by a bibliography of Artalejo (1999).

The batch Markov arrival process (BMAP) is a useful mathematical model for describing bursty traffic in modern communication networks and others. Readers may refer to recent publications for details, among which are Ramaswami (1980), Chapter 5 in Neuts (1989), Lucantoni (1991, 1993), Lucantoni and Neuts (1994), Lucantoni, Choudhury and Whitt (1994), Neuts (1995), Chakravarthy (2000) and Ferng and Chang (2001). In particular, Lee and Jeon (1999) and Lee (2000) used the supplementary variable method, which combined with the embedded Markov chain, to analyze the *BMAP/G/1* queues with finite or infinite waiting room.

The censoring technique has been successfully applied to discrete-time block-structured Markov chains and block-structured Markov renewal processes. Examples include Grassmann and Heyman (1990), Latouche and Ramaswami (1999), Zhao (2000), Li and Zhao (2002), Li and Zhao (2004) and Li and Zhao (2003). Dudin and Klimenok (2000) used the censoring technique to provide an approximate algorithm for calculating the stationary probability vector of an asymptotically quasi-Toeplitz 2-dimensional Markov chain. This paper applies the censoring technique to study a continuous-time level-dependent Markov chain of $M/G/1$ type and level-dependent Markov renewal processes of $M/G/1$ type, some new results are listed due to necessity for more applications.

Retrial queues have been studied by some authors in terms of the matrix-analytic method. Readers may refer to Neuts and Rao (1990), Diamond and Alfa (1995, 1998, 1999), Choi, Yang and Kim (1999), He, Li and Zhao (2000), Dudin and Klimenok (1999, 2000), Choi, Chung and Dudin (2001), Breuer, Dudin and Klimenok (2002), Chakravarthy and Dudin (2002, 2003) among others.

Queues with servers subject to breakdowns and repairs are often encountered in many practical applications such as in computer, manufacturing systems and communication networks. Because system performance deteriorates seriously by server breakdowns and the limitation of repair capacity, the study of queues with server breakdowns and repairs is not only important for theoretic investigations but also necessary for engineering applications. Such systems have been considered by some authors both from queueing viewpoint and from reliability viewpoint. Examples include the following four classes of important results in the literature. (i) To calculate queueing indices, some researchers proposed the so called *generalized service time* for unreliable queueing systems. Using the generalized service time, a unreliable queueing system can be simplified to an equivalent model, which is a reliable and ordinary queue. Therefore, the queueing indices, e.g., queue length, waiting time and busy period, are directly derived in terms of those results given in a corresponding ordinary queue with the generalized service time. Readers may further refer to Mitrany and Avi-Ttzhak (1968), Neuts and Lucantoni (1979) and Kulkarni and Choi (1990) among others. (ii) For the generalized service time, it has been illustrated that the life time of the server is a most crucial factor. When the life time of the server is exponential, the distribution of the generalized service time was given in Cao and Chen (1982) and Gnedenko and Kovalenko (1989). When the life time of the server is non-exponential and phase type, the distribution of the generalized service time was provided in Li (1996) for an $M/SM(PH/SM)/1$ repairable queue and Li, Tan and Sun (1999) for a $SM/PH(PH/PH)/1$ repairable queue. (iii) Reliability analysis of the server is very important in the study of unreliable queueing systems. Since the reliability indices are irrelevant to the computations of the queueing indices, the generalized service time is not helpful for computing the reliability indices. The stationary availability and the stationary failure frequency were given in Cao and Chen (1982), Li (1996), Li, Shi and Chao (1997), Li, Tan and Sun (1999) and Wang, Cao and Li (2001). (iv) For a block-structured repairable queue, Li (1997), Li (1996), Li, Tan and Sun (1999) and Li and Cao (2000) applied the matrix-analytic method to derive the stationary availability and failure frequency. On the other hand, Li (1997) and Hsu, Yuan and Li (2000) provided a uniformly convergent algorithm for computing the reliability function. This paper, to our best knowledge, first provides an analytic expression for the Laplace transform of the reliability function by means of the *RG*-factorization of a level-dependent Markov renewal processes of *M/G/1* type.

Retrial queues with unreliable servers have been discussed by some researchers, among whom are Kulkarni and Choi (1990), Yang and Li (1994), Artalejo (1994), Aissani (1994), Aissani and Artalejo (1998), Artalejo and Gómez-Corral (1998) and Wang, Cao and Li (2001). In this paper, a *BMAP/G/1* retrial queue with server breakdowns and repairers is considered both from the queueing viewpoint and from the reliability viewpoint, which is a more general class of retrial queues than those in literature.

The purpose of this paper is twofold. The first one is to provide expressions for both the reliability indices of the server and the queueing indices of the system. The expressions can be grouped into two classes as follows: i) Expressions based on the stationary probability vector of the system, for example, the stationary availability, failure frequency and queue length, and ii) expressions based on the Laplace transform of the distribution of a first passage time, for example, the reliability function and the busy period.

The other purpose is to extend the supplementary variable method (see Lee and Jeon (1999), Lee (2000)) to study more general block-structured queueing systems including the *BMAP/G/1* retrial queue. The key of successfully using the supplementary variable method is the new treatment suggested in this paper for boundary conditions of the system of differential equations for the model. The generalized approach reveals the underlying block structure by relating boundary conditions to a continuous-time level-dependent Markov chain of $M/G/1$ type or a level-dependent Markov renewal process of $M/G/1$ type. Therefore, the matrix-analytic method and the censoring technique can be fully utilized in the analysis and the *RG*-factorization of a continuous-time level-dependent Markov chain of $M/G/1$ type or a level-dependent Markov renewal process of $M/G/1$ type can then be established for a boundary solution.

The remainder of this paper is organized as follows. In the next section, we give a modeling description on the retrial *BMAP/G/1* queue with server breakdowns and repairs. In Section 3, we set up the system of stationary differential equations. In Section 4, we apply the censoring technique to solve the system of stationary differential equations, where the RG-factorization is a key. In Section 5, we derive three reliability indices of the server: The stationary availability, the stationary failure frequency and the reliability function. In Section 6, we express two queueing indices: The stationary queue length and the busy period. In Section 7, we propose two approximate algorithms and analyze their computational complexity. In Section 8, we give some numerical examples. Some concluding remarks are given in Section 9.

2. Model description

In this section, we describe the *BMAP/G/1/1* retrial queue with a server subject to breakdowns and repairs, which will be analyzed in this paper. The model is described as follows.

2.1. The arrival process

The arrivals to the retrial queue are modelled by a BMAP with m phases described by coefficient matrix sequence $\{D_k, k \geq 0\}$. The matrix D_0 has strictly negative diagonal entries and nonnegative off-diagonal entries, and is invertible. For $k \geq 1$, $D_k \geq 0$, we assume that $\sum_{k=1}^{\infty} kD_k$ is finite, and $D = \sum_{k=0}^{\infty} D_k$ is an irreducible infinitesimal generator with $De = 0$, where e is a column vector of ones. Let σ be the stationary probability vector of D. Then $\lambda = \sigma \sum_{k=1}^{\infty} kD_k e$ is the stationary arrival rate.

2.2. The service times

The service times $\{\chi_n, n \geq 1\}$ of the customers are assume to be i.i.d. random variables. The distribution function of the service time is expressed by

$$B(t) = P\{\chi_n \leq t\} = 1 - \exp\left\{-\int_0^t \mu(v)\, dv\right\}.$$

We assume that $E[\chi_n] = 1/\mu < +\infty$.

2.3. The life time and the repair time

The life time X of the server is exponential with mean life time $1/\alpha$ and does not change during the idle period of the server. The repair time Y of the server has the distribution function

$$V(y) = P\{Y \leq y\} = 1 - \exp\left\{-\int_0^y \beta(v)\, dv\right\}$$

with $E[Y] = 1/\beta < +\infty$.

2.4. The retrial rule

We assume that there is no waiting space in the retrial queue and the size of the orbit is infinite. If an arrival, either a primary or a retrial customer finds that there is no customer in the server, then it enters the server immediately and receives service, otherwise it enters the orbit and makes a retrial at a later time. Returning customers behave independently of each other and are persistent in the sense that they keep making retrials until they receive their requested service. Successive inter-retrial times $\{\xi_k, k \geq 1\}$ are i.i.d. exponentially distributed random variables with mean inter-retrial time $1/\theta$.

2.5. The service discipline

If the server is busy at the arrival epoch, then all these calls join the orbit, whereas if the server is free, then one of the arriving customers begins its service and the other calls join the orbit.

2.6. The repair discipline

When the server fails, it enters the state of failure and undergoes repair immediately. The customer who has been partially served has to wait to continue its service. As soon as the repair of the server is completed, the server enters the working state immediately and continues to serve the customer. We assume that the repaired server is as good as new and the service time is cumulative.

2.7. The independence

We assume that all the random variables defined above are independent.

Remark 1. Kulkarni and Choi (1990) considered a more general assumption on breakdowns: The server may be subject to either an active breakdown in the busy period or a passive breakdown in the idle period. As shown in Sections 2 and 3, the passive breakdowns only change those finitely-many equations corresponding to the idle period. Therefore, the passive breakdowns do not increase any difficulty in the analysis of such retrial queueing models with an unreliable server. For simplicity, in this paper we consider a model only with the active breakdowns.

3. The system of differential equations

In this section, we introduce several supplementary variables to make the model Markovian and set up the system of stationary differential equations for the model.

Let $\tilde{\chi}_n$ be the generalized service time of the nth customer, which is the length of time since the beginning of the service for the nth customer until the completion of the service. Clearly, $\tilde{\chi}_n$ includes the down time of the server due to server failures during the service period of the nth customer. It is easy to see, for example from Cao and Chen (1982), that the sequence $\tilde{\chi}_n$ for $n \geq 1$ are i.i.d. random variables and $E[\tilde{\chi}_n] = \frac{1}{\mu}(1 + \frac{\alpha}{\beta})$. It follows from Theorem 3 of Dudin and Klimenok (2000) or Liang and Kulkarni (1993) that if $\rho = \lambda E[\tilde{\chi}_n] = \frac{\lambda}{\mu}(1 + \frac{\alpha}{\beta}) < 1$, then the queueing system is stable. In the rest of this paper, we consider a stable system.

For the repairable *BMAP/G/1* retrial queue defined above, we denote by $N(t)$ the number of calls in the orbit at time t, and define the states of the server as

$$L(t) = \begin{cases} I, & \text{if the server is idle at time } t, \\ W, & \text{if the server is working at time } t, \\ R, & \text{if the server is under repair at time } t. \end{cases}$$

We introduce three variables $J(t)$, $S(t)$ and $R(t)$ representing the phase of the arrival process, the elapsed service time and the elapsed repair time at time t, respectively. Then $\{(L(t), N(t), J(t), S(t), R(t)): t \geq 0\}$ is a Markov process with state space expressed as

$$\Omega = \{(I, k, j) : k \geq 0, 1 \leq j \leq m\} \cup \{(W, k, j, x) : k \geq 0, 1 \leq j \leq m, x \geq 0\}$$
$$\cup \{(R, k, j, x, y) : k \geq 0, 1 \leq j \leq m, x \geq 0, \ y \geq 0\},$$

where k, j, x and y denotes the number of customers in the orbit, the phase of the arrival process, the amount of the elapsed service time and the elapsed repair time, respectively.

For $k \geq 0$, we define

$$P_{(I,k,j)}(t) = P\{L(t) = I, N(t) = k, J(t) = j\},$$

$$P_{(W,k,j)}(t,x)\,dx = P\{L(t) = W, N(t) = 1+k, J(t) = j, x \leq S(t) < x + dx\}$$

and

$$\begin{aligned} P_{(R,k,j)}(t,x,y)\,dy = P\{&L(t) = R, N(t) = 1+k, J(t) = j, S(t) \\ &= x, y \leq R(t) < y + dy\}. \end{aligned}$$

Write the above probabilities into vector form as

$$P_{I,k}(t) = \left(P_{(I,k,1)}(t), P_{(I,k,2)}(t), \ldots, P_{(I,k,m)}(t)\right),$$

$$P_{W,k}(t,x) = \left(P_{(W,k,1)}(t,x), P_{(W,k,2)}(t,x), \ldots, P_{(W,k,m)}(t,x)\right)$$

and

$$P_{R,k}(t,x,y) = \left(P_{(R,k,1)}(t,x,y), P_{(R,k,2)}(t,x,y), \ldots, P_{(R,k,m)}(t,x,y)\right).$$

Since we are interested in the stationary behavior of the system, define

$$P_{I,k} = \lim_{t\to+\infty} P_{I,k}(t),\ P_{W,k}(x) = \lim_{t\to+\infty} P_{W,k}(t,x),\ P_{R,k}(x,y) = \lim_{t\to+\infty} P_{R,k}(t,x,y).$$

The joint probability density $\left\{P_{I,k}, P_{W,k}(x), P_{R,k}(x,y), k \geq 0\right\}$ satisfies the following system of differential equations:

$$\frac{d}{dx}P_{W,0}(x) = P_{W,0}(x)\{D_0 - [\alpha + \mu(x)]I\} + \int_0^{+\infty} \beta(y) P_{R,0}(x,y)\,dy, \tag{1}$$

$$\begin{aligned} \frac{d}{dx}P_{W,k}(x) = {} & P_{W,k}(x)\{D_0 - [\alpha + \mu(x)]I\} + \sum_{i=0}^{k-1} P_{W,i}(x) D_{k-i} \\ & + \int_0^{+\infty} \beta(y) P_{R,k}(x,y)\,dy, \qquad k \geq 1, \end{aligned} \tag{2}$$

$$\frac{\partial}{\partial y}P_{R,0}(x,y) = P_{R,0}(x,y)[D_0 - \beta(y)I], \tag{3}$$

$$\frac{\partial}{\partial y}P_{R,k}(x,y) = P_{R,k}(x,y)[D_0 - \beta(y)I] + \sum_{i=0}^{k-1} P_{R,i}(x,y) D_{k-i}, \quad k \geq 1, \tag{4}$$

with the boundary conditions

$$P_{I,k}\,(k\theta I - D_0) = \int_0^{+\infty} \mu\,(x)\, P_{W,k}\,(x)\, dx, \qquad k \geq 0, \tag{5}$$

$$P_{W,k}\,(0) = \sum_{i=0}^{k} P_{I,i}\, D_{k+1-i} + (k+1)\,\theta\, P_{I,k+1}, \qquad k \geq 0, \tag{6}$$

$$P_{R,k}\,(x, 0) = \alpha\, P_{W,k}\,(x)\,, \qquad k \geq 0, \tag{7}$$

and the normalization condition

$$\sum_{k=0}^{\infty} \left[P_{I,k} + \int_0^{+\infty} P_{W,k}\,(x)\, dx + \int_0^{+\infty} \int_0^{+\infty} P_{R,k}\,(x, y)\, dx dy \right] e = 1. \tag{8}$$

4. Solving the system of differential equations

In this section, we provide an approach for solving the equations (1) to (8). There are two crucial steps: We first express $P_{W,k}\,(x)$ and $P_{R,k}\,(x, y)$ in terms of boundary probabilistic vectors $P_{W,k}\,(0)$ by recognizing a new BMAP, and then provide a method for obtaining $P_{I,k}$ and $P_{W,k}\,(0)$ by converting boundary equations into the stationary equations of a level-dependent Markov chain of $M/G/1$ type. This solution will be used to express interesting performance measures of the system in later sections.

Let

$$D^*\,(z) = \sum_{k=0}^{\infty} z^k D_k, \quad \tilde{v}(D^*\,(z)) = \int_0^{+\infty} \exp\left\{D^*\,(z)\, y\right\} dV\,(y)\,.$$

The following lemma recognizes a new BMAP, which appears in the process of solving the system of differential equations.

Lemma 1. *Let*

$$\Psi^*(z) = \sum_{k=0}^{\infty} z^k \Psi_k = D^*(z) - \alpha[I - \tilde{v}(D^*(z))]. \tag{9}$$

Then Ψ_k for $k \geq 0$ are coefficient matrices of a BMAP of size m.

Proof: To prove this lemma, we need to show that the following three conditions are satisfied: (i) The diagonal entries of Ψ_0 are strictly negative, the off-diagonal entries are nonnegative, and Ψ_0 is invertible. (ii) For $k \geq 1$, $\Psi_k \geq 0$ and $\sum_{k=1}^{\infty} k\Psi_k < +\infty$. iii) $\Psi = \sum_{k=0}^{\infty} \Psi_k$ is irreducible and $\Psi e = 0$.

(i) It follows from (9) that

$$\Psi_0 = D_0 - \alpha\,[I - \tilde{v}\,(D_0)]\,. \tag{10}$$

It is clear from (10) that the off-diagonal entries of Ψ_0 are nonnegative, since $\tilde{v}(D_0) \geq 0$. Noting that the ith diagonal entry of the matrix $\tilde{v}(D_0)$ is the conditional probability that the BMAP returns to state i and no arrival occurs during a repair time, given that the BMAP starts in state i, we obtain that the ith diagonal entry of the matrix $I - \tilde{v}(D_0)$ is nonnegative. Hence, the diagonal entries of Ψ_0 are strictly negative according to the assumption of D_0. Furthermore, the real parts of the eigenvalues of Ψ_0 are all strictly negative according to Gerŝgorin Theorem (see Horn and Johnson (1985)), and so Ψ_0 is invertible.

(ii) It is clear that for $k \geq 1$, $\frac{d^k}{dz^k}\left[\tilde{v}(D^*(z))\right]_{|z=0} \geq 0$. Since

$$\Psi_k = D_k + \alpha \frac{1}{k!}\frac{d^k}{dz^k}[\tilde{v}(D^*(z))]_{|z=0},$$

$D_k \geq 0$ for $k \geq 1$ and $\alpha > 0$, we obtain that for $k \geq 1$, $\Psi_k \geq 0$. Using $\sum_{k=1}^{\infty} kD_k < +\infty$, we obtain

$$\sum_{k=1}^{\infty} k\Psi_k = \left[I + \alpha \int_0^{+\infty} x \exp\{Dx\}\, dV(x)\right] \sum_{k=1}^{\infty} kD_k < +\infty.$$

(iii) Noting that

$$\sum_{k=0}^{\infty} \Psi_k = D + \alpha\tilde{v}(D) - \alpha I,$$

it is clear that Ψ is irreducible, since D is irreducible and $\tilde{v}(D) \geq 0$. Noting that $De = 0$ and $[I - \tilde{v}(D)]\, e = 0$, it is obvious that $\Psi e = 0$. This completes the proof. □

Remark 2. The BMAP with coefficient matrix sequence $\{\Psi_k\}$ may be regarded as a generalized arrival process, which is composed of the sum of two parts: the first is the original BMAP with coefficient matrix sequence $\{D_k\}$ while the other is an additional BMAP with coefficient matrix sequence

$$\left\{-\alpha\left[I - \tilde{v}(D_0)\right], \alpha\frac{1}{k!}\frac{d^k}{dz^k}[\tilde{v}(D^*(z))]_{|z=0}, k = 1, 2, \ldots\right\}.$$

The additional BMAP is due to the server subject to breakdowns and repairs. Note that Lemma 1 is crucial in computations of the reliability indices as demonstrated in Section 5.

For the two BMAPs having coefficient matrix sequences $\{D_k\}$ and $\{\Psi_k\}$, let $K^D(t)$ and $K^\Psi(t)$ denote the numbers of arrivals in the time interval $[0, t)$, respectively, and $J^D(t)$ and $J^\Psi(t)$ the phases at time t, respectively. We introduce the conditional probabilities for the two BMAPs by

$$P_{j,j'}^D(n,t) = P\{K^D(t) = n, J^D(t) = j' \mid K^D(0) = 0, J^D(0) = j\},$$

and

$$P_{j,j'}^\Psi(n,t) = P\left\{K^\Psi(t) = n, J^\Psi(t) = j' \mid K^\Psi(0) = 0, J^\Psi(0) = j\right\}.$$

Let $P^D(n,t)$ and $P^\Psi(n,t)$ be the matrices with entries $P^D_{j,j'}(n,t)$ and $P^\Psi_{j,j'}(n,t)$ for $1 \le j$, $j' \le m$, respectively. Then it follows from Neuts (1989) or Lucantoni (1991) that

$$P^*_D(z,t) = \sum_{n=0}^{\infty} z^n P^D(n,t) = \exp\{D^*(z)t\} \tag{11}$$

and

$$P^*_\Psi(z,t) = \sum_{n=0}^{\infty} z^n P^\Psi(n,t) = \exp\{\Psi^*(z)t\}. \tag{12}$$

We write

$$\bar{B}(x) = 1 - B(x), \bar{V}(y) = 1 - V(y),$$
$$P^*_W(z,x) = \sum_{k=0}^{\infty} z^k P_{W,k}(x), P^*_R(z,x,y) = \sum_{k=0}^{\infty} z^k P_{R,k}(x,y).$$

Now, we solve the system of matrix equations (1) to (8). It follows from (3) and (4) that

$$\frac{\partial}{\partial y} P^*_R(z,x,y) = P^*_R(z,x,y)[D^*(z) - \beta(y)I],$$

hence from (7) we obtain

$$\begin{aligned} P^*_R(z,x,y) &= P^*_R(z,x,0)\exp\{D^*(z)y\}\bar{V}(y) \\ &= \alpha P^*_W(z,x)\exp\{D^*(z)y\}\bar{V}(y). \end{aligned} \tag{13}$$

It follows from (1) and (2), together with (13), that

$$\frac{\partial}{\partial x} P^*_W(z,x) = P^*_W(z,x)\{D^*(z) - \alpha[I - \tilde{v}(D^*(z))] - \mu(x)I\}.$$

Hence,

$$\begin{aligned} P^*_W(z,x) &= P^*_W(z,0)\exp\{\{D^*(z) - \alpha[I - \tilde{v}(D^*(z))]\}x\}\bar{B}(x) \\ &= P^*_W(z,0)\exp\{\Psi^*(z)x\}\bar{B}(x), \end{aligned} \tag{14}$$

which, together with (12), leads to

$$P_{W,k}(x) = \sum_{i=0}^{k} P_{W,i}(0) P^\Psi(k-i,x)\bar{B}(x). \tag{15}$$

Similarly, (13), together with (11) and (15), leads to

$$P_{R,l}(x, y) = \alpha \sum_{k=0}^{l} P_{W,k}(x) P^D(l-k, y) \bar{V}(y)$$

$$= \alpha \sum_{k=0}^{l} \sum_{i=0}^{k} P_{W,i}(0) P^\Psi(k-i, x) P^D(l-k, y) \bar{B}(x) \bar{V}(y). \tag{16}$$

Equations (15) and (16) provide a solution for $P_{W,k}(x)$ and $P_{R,k}(x, y)$ in terms of $P_{W,k}(0)$, $k \geq 0$. In order to completely solve the system of differential equations, we still need to determine the vectors $P_{W,k}(0)$ and $P_{I,k}$ for $k \geq 0$ from the boundary equations (5) and (6), and the normalization condition (8). Define

$$P_{IW} = \left(P_{I,0}, P_{W,0}(0), P_{I,1}, P_{W,1}(0), P_{I,2}, P_{W,2}(0), P_{I,3}, P_{W,3}(0), \ldots\right),$$

$$C_k = \int_0^{+\infty} P^\Psi(k, x)\, dB(x), \; k \geq 0,$$

$$A_0^{(k)} = \begin{pmatrix} 0 & k\theta I \\ 0 & 0 \end{pmatrix}, \; k \geq 1, \tag{17}$$

$$A_1^{(k)} = \begin{pmatrix} -k\theta I + D_0 & D_1 \\ C_0 & -I \end{pmatrix}, \; k \geq 0, \tag{18}$$

$$A_k = \begin{pmatrix} 0 & D_k \\ C_{k-1} & 0 \end{pmatrix}, \; k \geq 2, \tag{19}$$

and

$$Q = \begin{pmatrix} A_1^{(0)} & A_2 & A_3 & A_4 & \cdots \\ A_0^{(1)} & A_1^{(1)} & A_2 & A_3 & \cdots \\ & A_0^{(2)} & A_1^{(2)} & A_2 & \cdots \\ & & A_0^{(3)} & A_1^{(3)} & \cdots \\ & & & \ddots & \ddots \end{pmatrix}. \tag{20}$$

According to the above definitions and the expression for $P_{W,k}(x)$ in (15), the boundary equations (5) and (6) can be written as

$$P_{IW} Q = 0. \tag{21}$$

In what follows we show that the matrix Q is the infinitesimal generator of a continuous-time positive recurrent Markov chain. Therefore, the unique stationary probability vector X of Q can be used to determine the vectors $P_{W,k}(0)$ and $P_{I,k}$ for $k \geq 0$. It is clear that $P_{IW} = \gamma X$. Let $X = (\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots)$, where $\mathbf{x}_k = \left(\mathbf{x}_{k,1}, \mathbf{x}_{k,2}\right)$ for $k \geq 0$ and the size of each vector $\mathbf{x}_{k,j}$, $k \geq 0$ and $j = 1, 2$, is m. The normalization condition (8) and the expressions for $P_{W,k}(x)$ in (15) and $P_{R,k}(x, y)$ in (16) lead to

$$\gamma = \frac{1}{\sum_{k=0}^{\infty}\sum_{i=0}^{k}\left(\mathbf{x}_{i,1} + \mathbf{x}_{i,2}F_{k-i}\right)e + \alpha\sum_{l=0}^{\infty}\sum_{k=0}^{l}\sum_{i=0}^{k}\mathbf{x}_{i,2}F_{k-i}H_{l-k}}, \tag{22}$$

where

$$F_k = \int_0^{+\infty} P^{\Psi}(k, x)\,\bar{B}(x)\,dx,\;\; H_k = \int_0^{+\infty} P^{D}(k, y)\,\bar{V}(y)\,dy. \tag{23}$$

Theorem 1. *The matrix Q is the infinitesimal generator of a continuous-time irreducible positive recurrent Markov chain.*

Proof: From the definition of Q, it can be easily verified that the matrix Q can be served as the infinitesimal generator of a continuous-time irreducible Markov chain. In what follows we only need to prove that it is positive recurrent.

For each $k \geq 1$,

$$A_0^{(k)} + A_1^{(k)} + \sum_{l=2}^{\infty} A_l = \begin{pmatrix} D_0 - k\theta I & D_+ + k\theta I \\ C & -I \end{pmatrix},$$

where $D_+ = \sum_{l=1}^{\infty} D_l$ and $C = \sum_{l=0}^{\infty} C_l$. It is clear that the transition rate matrix $A_0^{(k)} + A_1^{(k)} + \sum_{l=2}^{\infty} A_l$ is irreducible and positive recurrent. Let $(y_1^{(k)}, y_2^{(k)})$ be the stationary probability vector of $A_0^{(k)} + A_1^{(k)} + \sum_{l=2}^{\infty} A_l$. Then

$$\left(y_1^{(k)}, y_2^{(k)}\right)\begin{pmatrix} D_0 - k\theta I & D_+ + k\theta I \\ C & -I \end{pmatrix} = 0,$$

hence, solving this equation gives

$$y_1^{(k)}\left[\frac{1}{k\theta}(D_0 + D_+C) + C - I\right] = 0.$$

Noting that the matrix C is irreducible and stochastic, and the matrix $D_0 + D_+$ is an infinitesimal generator, it is clear that $\frac{1}{k\theta}(D_0 + D_+C) + C - I$ is an irreducible infinitesimal generator of size m for each $k \geq 1$. Thus, for each $k \geq 1$ the Markov chain $\frac{1}{k\theta}(D_0 + D_+C) + C - I$ is positive recurrent. Let $w^{(k)}$ be the stationary probability vector of $\frac{1}{k\theta}(D_0 + D_+C) + C - I$. Then

$$y_1^{(k)} = \frac{w^{(k)}}{1 + k\theta + w^{(k)}D_+e} \quad \text{and} \quad y_2^{(k)} = \frac{w^{(k)}(D_+ + k\theta I)}{1 + k\theta + w^{(k)}D_+e},$$

Noting that, as $k \to \infty$,

$$\frac{1}{k\theta}(D_0 + D_+C) + C - I \to C - I,$$

it is clear that $w^{(k)} \to w$, as $k \to \infty$, where w is the stationary probability vector of the irreducible infinitesimal generator $C - I$. Thus, as $k \to \infty$,

$$y_1^{(k)} = \frac{w^{(k)}}{1 + k\theta + w^{(k)} D_+ e} \to 0 \quad \text{and} \quad y_2^{(k)} = \frac{w^{(k)}\left(D_+ + k\theta I\right)}{1 + k\theta + w^{(k)} D_+ e} \to w.$$

As $k \to \infty$, some simple computations lead to

$$\left(y_1^{(k)}, y_2^{(k)}\right) A_0^{(k)} e = k\theta y_1^{(k)} e = \frac{w^{(k)} C \left(I - \frac{1}{k\theta} D_0\right)^{-1} e}{1 + \frac{1}{k\theta} w^{(k)} C \left(I - \frac{1}{k\theta} D_0\right)^{-1} e} \to 1 \tag{24}$$

and

$$\begin{aligned} \left(y_1^{(k)}, y_2^{(k)}\right) \sum_{l=2}^{\infty} (l-1) A_l e \;&=\; y_1^{(k)} \sum_{l=2}^{\infty} (l-1) D_l e + y_2^{(k)} \sum_{l=2}^{\infty} (l-1) C_{l-1} e \\ &\to w \sum_{l=2}^{\infty} (l-1) C_{l-1} e = \frac{\lambda}{\mu} < 1 \end{aligned} \tag{25}$$

due to the stable condition $\rho = \frac{\lambda}{\mu}(1 + \frac{\alpha}{\beta}) < 1$. It follows from (24) and (25) that

$$\lim_{k\to\infty} \left(y_1^{(k)}, y_2^{(k)}\right) A_0^{(k)} e > \lim_{k\to\infty} \left(y_1^{(k)}, y_2^{(k)}\right) \sum_{l=2}^{\infty} (l-1) A_l e. \tag{26}$$

Thus, there always exists a positive integer N Large enough such that for all $k > N$,

$$\left(y_1^{(k)}, y_2^{(k)}\right) A_0^{(k)} e > \left(y_1^{(k)}, y_2^{(k)}\right) \sum_{l=2}^{\infty} (l-1) A_l e. \tag{27}$$

It is easy to check that

$$\sum_{l=2}^{\infty} (l-1) A_l e < +\infty. \tag{28}$$

Therefore, it is easy to see from (27) and (28) that the continuous-time irreducible Markov chain Q is positive recurrent based on the principle of mean drift (for example, Proposition 4.6 in Asmussen (1987) for discrete-time case). □

What is left now is to efficiently determine the stationary probability vector X of Q. The procedure here is based on the censoring technique and a RG-factorization. To do this, we write

$$Q_k = \begin{pmatrix} A_1^{(k)} & A_2 & A_3 & \cdots \\ A_0^{(k+1)} & A_1^{(k+1)} & A_2 & \cdots \\ & A_0^{(k+2)} & A_1^{(k+2)} & \cdots \\ & & \ddots & \ddots \end{pmatrix}, \quad k \geq 1.$$

For Q_k we denote by $(\hat{Q}_{1,1}^{(k)\,T}, \hat{Q}_{2,1}^{(k)\,T}, \ldots)^T$ the first block-column of its fundamental matrix $\hat{Q}_k = -Q_k^{-1}$, which is the minimal nonnegative inverse of $-Q_k$. We define

$$R_j^{(k)} = \sum_{i=1}^{\infty} A_{i+j}\,\hat{Q}_{i,1}^{(k+1)}, \quad k \geq 0,\ j \geq 1, \tag{29}$$

$$G^{(k)} = \hat{Q}_{1,1}^{(k)} A_0^{(k)}, \quad k \geq 1, \tag{30}$$

and

$$U_k = A_1^{(k)} + R_1^{(k)} A_0^{(k+1)}, \quad k \geq 0. \tag{31}$$

Based on the censoring technique, or a similar argument employed in Lemma 4 of Li and Zhao (2002), we can have the following Lemma.

Lemma 2. For $k \geq 1$ and $j \geq 2$,

$$\hat{Q}_{j,1}^{(k)} = G^{(k+j-1)} G^{(k+j-2)} \cdots G^{(k+1)} \left(-U_k^{-1}\right),$$

and

$$U_k = -\left[\hat{Q}_{1,1}^{(k)}\right]^{-1}.$$

Lemma 3.

(i) *For $k \geq 0$ and $l \geq 1$,*

$$R_l^{(k)} = \left[A_{l+1} + A_{l+2} G^{(k+2)} + A_{l+3} G^{(k+3)} G^{(k+2)} + \cdots\right] \left(-U_{k+1}^{-1}\right). \tag{32}$$

(ii) *For $k \geq 0$,*

$$U_k = A_1^{(k)} + A_2 G^{(k+1)} + A_3 G^{(k+2)} G^{(k+1)} + A_4 G^{(k+3)} G^{(k+2)} G^{(k+1)} + \cdots. \tag{33}$$

(iii) *The matrix sequence $\{G^{(k)}\}$ is the minimal nonnegative solution to the system of matrix equations*

$$A_0^{(k)} + A_1^{(k)} G^{(k)} + A_2 G^{(k+1)} G^{(k)} + A_3 G^{(k+2)} G^{(k+1)} G^{(k)} + \cdots = 0,\ k \geq 1. \tag{34}$$

Proof: (i) It follows from Lemma 2 that

$$\hat{Q}_{1,1}^{(k+1)} = -U_{k+1}^{-1}$$

and

$$\hat{Q}_{j,1}^{(k+1)} = G^{(k+j)} G^{(k+j-1)} \cdots G^{(k+2)} \left(-U_{k+1}^{-1}\right), \ j \geq 2.$$

Thus, (29) becomes

$$R_l^{(k)} = \left[A_{l+1} + A_{l+2} G^{(k+2)} + A_{l+3} G^{(k+3)} G^{(k+2)} + \cdots\right] \left(-U_{k+1}^{-1}\right).$$

(ii) It follows from (31) that

$$U_k = A_1^{(k)} + A_2 G^{(k+1)} + A_3 G^{(k+2)} G^{(k+1)} + A_4 G^{(k+3)} G^{(k+2)} G^{(k+1)} + \cdots.$$

(iii) Noting that $(-U_k)\left(-U_k^{-1}\right) = I$ and $\left(-U_k^{-1}\right) A_0^{(k)} = G^{(k)}$, we obtain $(-U_k)\, G^{(k)} = A_0^{(k)}$, which is equivalent to

$$A_0^{(k)} + A_1^{(k)} G^{(k)} + A_2 G^{(k+1)} G^{(k)} + A_3 G^{(k+2)} G^{(k+1)} G^{(k)} + \cdots = 0.$$

A similar discussion to the proof of Lemma 1.2.3 in Neuts (1989) leads to the conclusion that $\{G^{(k)}\}$ is the minimal nonnegative solution to the system of matrix equations in (34). This completes the proof. □

Remark 3. In principle, the matrix sequence $\{G^{(k)}\}$ can be numerically computed. Once $\{G^{(k)}\}$ is given, we can obtain the R-measure $\{R_l^{(k)}\}$ and the matrix sequence $\{U_k\}$ according to Lemma 3. Readers may refer to Dudin and Klimenok (1999, 2000).

Theorem 2.

$$Q = (I - R_U)\, U_D\, (I - G_L), \tag{35}$$

where

$$(I - R_U) = \begin{pmatrix} I & -R_1^{(0)} & -R_2^{(0)} & -R_3^{(0)} & \cdots \\ & I & -R_1^{(1)} & -R_2^{(1)} & \cdots \\ & & I & -R_1^{(2)} & \cdots \\ & & & I & \cdots \\ & & & & \ddots \end{pmatrix},$$

U_D is the diagonal matrix in block form with the diagonal entries equal to U_k for $k \geq 0$, or $U_D = diag\,(U_0, U_1, U_2, \ldots)$, and

$$(I - G_L) = \begin{pmatrix} I & & & & \\ -G^{(1)} & I & & & \\ & -G^{(2)} & I & & \\ & & -G^{(3)} & I & \\ & & & \ddots & \ddots \end{pmatrix}.$$

Proof: To prove (35), we first compute the right hand side of (35), and then compare it with the left hand side. For the right hand side of (35), a) the entries of the first block-row are $U_0 + R_1^{(0)} U_1 G^{(1)}$ and $-R_l^{(0)} U_l + R_{l+1}^{(0)} U_{l+1} G^{(l+1)}$, $l \geq 1$; and b) the entries of the kth block-row are $-U_k G^{(k)}$, $U_k + R_1^{(k)} U_{k+1} G^{(k+1)}$ and $-R_l^{(k)} U_{k+1} + R_{l+1}^{(k+1)} U_{k+2} G^{(k+2)}$.

It follows from (32) that

$$R_1^{(0)} U_1 G^{(1)} = -\left[A_2 G^{(1)} + A_3 G^{(2)} G^{(1)} + A_4 G^{(3)} G^{(2)} G^{(1)} + \cdots\right]$$

and from (33) that

$$U_0 + R_1^{(0)} U_1 G^{(1)} = A_1^{(0)}.$$

It follows from (32) that

$$-R_l^{(0)} U_l = A_{l+1} + A_{l+2} G^{(l+1)} + A_{l+3} G^{(l+2)} G^{(l+1)} + A_{l+4} G^{(l+3)} G^{(l+2)} G^{(l+1)} + \cdots$$

and

$$R_{l+1}^{(0)} U_{l+1} G^{(l+1)} = -\left[A_{l+2} G^{(l+1)} + A_{l+3} G^{(l+2)} G^{(l+1)} + A_{l+4} G^{(l+3)} G^{(l+2)} G^{(l+1)} + \cdots\right].$$

Hence,

$$-R_l^{(0)} U_l + R_{l+1}^{(0)} U_{l+1} G^{(l+1)} = A_{l+1}.$$

Similarly, for the kth block-row we can obtain

$$\begin{aligned} -U_k G^{(k)} &= A_0^{(k)}, \\ U_k + R_1^{(k)} U_{k+1} G^{(k+1)} &= A_1^{(k)} \end{aligned}$$

and

$$-R_l^{(k)} U_{k+1} + R_{l+1}^{(k+1)} U_{k+2} G^{(k+2)} = A_{l+1}.$$

Therefore, the right hand side of (35) is the same as its left hand side. This completes the proof. □

The equation (35) is called the RG-factorization for continuous-time level-dependent Markov chains of $M/G/1$ type.

Remark 4. Zhao (2000) applied the censoring technique to provide the RG-factorization for an irreducible Markov chain of GI/G/1 type. While we use the same technique in this paper to derive the RG-factorization for an irreducible continuous-time level-dependent Markov chain of $M/G/1$ type, which is always very useful in modeling and analyzing some practical systems. As seen from those block-entries in (17) and (18), the continuous-time level-dependent Markov chain of $M/G/1$ type given in (20) can not be rewritten as a discrete-time Markov chain of $GI/G/1$ type as in Zhao (2000).

The following corollary expresses the stationary probability vector of Q in terms of the RG-factorization. The proof is clear according to a solving procedure provided in Subsection 5.1 of Li and Zhao (2002).

Corollary 3. *The stationary probability vector of Q is given by*

$$\begin{cases} \mathbf{x}_0 = \tau z_0, \\ \mathbf{x}_k = \sum_{i=0}^{k-1} \mathbf{x}_i R_{k-i}^{(i)}, \quad k \geq 1, \end{cases}$$

where z_0 is the stationary probability vector of the transition rate matrix U_0 and the scalar τ is uniquely determined by $\sum_{k=0}^{\infty} \mathbf{x}_k e = 1$.

The solution of the system of differential equations for the vectors $P_{I,k}$, $P_{W,k}(x)$ and $P_{R,k}(x, y)$ for $k \geq 0$ based on the above discussion is summarized into the following theorem.

Theorem 4. *If the system is stable, then for $k \geq 0$,*

$$\begin{cases} P_{I,k} = \gamma \mathbf{x}_{k,1}, \\ P_{W,k}(x) = \gamma \sum_{i=0}^{k} \mathbf{x}_{k,2} P^{\Psi}(k-i, x) \bar{B}(x), \\ P_{R,k}(x, y) = \alpha\gamma \sum_{l=0}^{k} \sum_{i=0}^{l} \mathbf{x}_{i,2} P^{\Psi}(l-i, x) P^{D}(k-l, y) \bar{B}(x) \bar{V}(y), \end{cases}$$

where γ and $\mathbf{x}_k = (\mathbf{x}_{k,1}, \mathbf{x}_{k,2})$ for $k \geq 0$ are given in (22) and Corollary 3, respectively.

Remark 5. The method proposed in this section can be used in principle to deal with a retrial queue with a more general total retrial rate, where the total retrial rate is a function $f(n, \theta)$, n is the number of customers in the orbit and θ is a parameter. For example, $f(n, \theta) = \sum_{i=0}^{M} a_i(\theta) n^i$ or $f(n, \theta) = C_1 \ln n + C_2 e^{n\theta}$. The case with linear retrial rate $f(n, \theta) = n\theta + \gamma$ was studied in Dudin and Klimenok (2000).

5. Reliability indices

In this section, we explicitly express three reliability indices of the server: The stationary availability, the stationary failure frequency and the reliability function.

To obtain the stationary availability and the stationary failure frequency of the server, we need the following lemma.

Lemma 4. *If the system is stable, then*

(i) *the probability that the server is idle is* $P_I = 1 - \frac{\lambda}{\mu}(1 + \frac{\alpha}{\beta})$,
(ii) *the probability that the server is working is* $P_W = \frac{\lambda}{\mu}$,
(iii) *the probability that the server is under repair is* $P_R = \frac{\lambda}{\mu}\frac{\alpha}{\beta}$.

Proof: It follows from (5) and (14) that

$$\theta z \frac{d}{dz} P_I^* (z) - P_I^* (z) D_0 = P_W^* (z, 0) \tilde{b} \left(\Psi^* (z) \right), \tag{36}$$

where

$$P_I^* (z) = \sum_{k=0}^{\infty} z^k P_{I,k}, \quad \tilde{b} \left(\Psi^* (z) \right) = \int_0^{+\infty} \exp \left\{ \Psi^* (z) x \right\} dB (x).$$

It follows from (6) that

$$P_W^* (z, 0) = \frac{1}{z} P_I^* (z) \left[D^* (z) - D_0 \right] + \theta \frac{d}{dz} P_I^* (z). \tag{37}$$

From (36) and (37) we obtain

$$P_W^* (z, 0) = P_I^* (z) D^* (z) \left[zI - \tilde{b} \left(\Psi^* (z) \right) \right]^{-1}.$$

Noting that

$$\int_0^{+\infty} \exp \left\{ \Psi^* (z) x \right\} \bar{B} (x) dx = \left[I - \tilde{b} \left(\Psi^* (z) \right) \right] \left[-\Psi^* (z) \right]^{-1}$$

and

$$\int_0^{+\infty} \exp \left\{ D^* (z) x \right\} \bar{V} (x) dx = \left[I - \tilde{v} \left(D^* (z) \right) \right] \left[-D^* (z) \right]^{-1},$$

using (13), (14) and (37) yields

$$\begin{aligned} P^* (z) &= P_I^* (z) + \int_0^{+\infty} P_W^* (z, x) dx + \int_0^{+\infty} \int_0^{+\infty} P_R^* (z, x, y) dx dy \\ &= P_I^* (z) \left\{ I + D^* (z) \left[zI - \tilde{b} \left(\Psi^* (z) \right) \right]^{-1} \left[I - \tilde{b} \left(\Psi^* (z) \right) \right] \left[-\Psi^* (z) \right]^{-1} \right. \\ &\quad \left. \times \left\{ I + \alpha \left[I - \tilde{v} \left(D^* (z) \right) \right] \left[-D^* (z) \right]^{-1} \right\} \right\}. \end{aligned} \tag{38}$$

For $z \geq 0$, we denote by $\chi(z)$ and $e(z)$ the eigenvalue with maximal real part of the matrix $D^*(z)$ and the associated right eigenvector with the first entry normalized to one, respectively. It is obvious that $\lim_{z\to 1}\chi(z) = 0$, $\lim_{z\to 1}\chi'(z) = \lambda$ and $\lim_{z\to 1} e(z) = e$. It follows from (38) that

$$P^*(z)\, e(z) = P_I^*(z)\, e(z)\, \frac{z-1}{z - \tilde{b}\left(\chi(z) - \alpha\left[1 - \tilde{v}(\chi(z))\right]\right)}. \tag{39}$$

Noting that, after some calculations,

$$\lim_{z\to 1} \frac{z-1}{z - \tilde{b}\left(\chi(z) - \alpha\left[1 - \tilde{v}(\chi(z))\right]\right)} = \frac{1}{1 - \frac{\lambda}{\mu}\left(1 + \frac{\alpha}{\beta}\right)}$$

and

$$\lim_{z\to 1} P^*(z)\, e(z) = 1,$$

we obtain

$$P_I = P_I^*(1)\, e(1) = 1 - \frac{\lambda}{\mu}\left(1 + \frac{\alpha}{\beta}\right).$$

We can similarly obtain that $P_W = \frac{\lambda}{\mu}$ and $P_R = \frac{\lambda}{\mu}\frac{\alpha}{\beta}$. This completes the proof. □

Remark 6. This lemma shows that we can determine the scalar function $P^*(z)\, e(z)$ by (36) to (39). However, we can not explicitly obtain the vector function $P^*(z)$. This is the main reason why it is necessary for us to provide an approach for solving the equations (1) to (8) in Section 4. Meanwhile, it is also easy to see the basic difficulty of using the standard method (e.g., see Subsection 1.2.2 in Falin and Templeton (1997)) to deal with the *BMAP/G/1* retrial queue and more generally, retrial queues of *M/G/1* type.

Let

$$A(t) = P\{\text{the server is up at time } t\}$$

and define the stationary availability of the server as $A = \lim_{t\to+\infty} A(t)$. We denote by W_f the stationary failure frequency of the server.

Theorem 5. *If the system is stable, then*

(i) *the stationary availability of the server is given by*

$$A = 1 - \frac{\lambda}{\mu}\frac{\alpha}{\beta};$$

(ii) *the stationary failure frequency of the server is given by*

$$W_f = \alpha\frac{\lambda}{\mu}.$$

Proof: Noting that

$$A = \sum_{k=0}^{\infty} \left[P_{I,k} + \int_0^{+\infty} P_{W,k}(x)\,dx \right] e = \left[P_I^*(1) + \int_0^{+\infty} P_W^*(1,x)\,dx \right] e = P_I + P_W$$

and

$$W_f = \alpha \left[\sum_{k=0}^{\infty} \int_0^{+\infty} P_{W,k}(x)\,dx \right] e = \alpha \int_0^{+\infty} P_W^*(1,x)\,dxe = \alpha P_W.$$

This completes the proof. □

Remark 7. It is easy to see from (13) to (16) that Lemma 1 is a key to express the vector generating function $P_W^*(z,x)$, which is necessary to derive the two reliability indices: The stationary availability and failure frequency, as shown in the proof of Theorem 5.

Now, we derive the reliability function of the server. Let ξ be the time to the first failure of the server. We write

$$\Re(t) = P\{\xi > t\},$$

which is called the reliability function of the server.

To derive the expression for $\Re(t)$, we first treat all the states of failure of the server as an absorbing state '$*$', we then obtain a transient Markov process $\{(L(t), N(t), J(t), S(t)): t \geq 0\}$ on state space

$$\Omega_0 = \{(I,k,i) : k \geq 0, 1 \leq i \leq m\} \cup \{(W,k,i,x) : k \geq 0, 1 \leq i \leq m, x \geq 0\} \cup \{*\}.$$

If we use the same notation in Section 2, then the systems of differential equations are given by

$$\frac{d}{dt} P_{I,k}(t) = P_{I,k}(t)[D_0 - k\theta I] + \int_0^{+\infty} \mu(x) P_{W,k}(t,x)\,dx, \ \ k \geq 0, \tag{40}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} \right) P_{W,0}(t,x) = P_{W,0}(t,x)\{D_0 - [\alpha + \mu(x)] I\}, \tag{41}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x} \right) P_{W,k}(t,x) = P_{W,k}(t,x)\{D_0 - [\alpha + \mu(x)] I\} + \sum_{i=0}^{k-1} P_{W,i}(t,x) D_{k-i}, \quad k \geq 1. \tag{42}$$

The boundary conditions are given by

$$P_{W,k}(t,0) = \sum_{i=0}^{k} P_{I,i}(t) D_{k+1-i} + (k+1)\theta P_{I,k+1}(t), \quad k \geq 0, \tag{43}$$

and the initial conditions are given by

$$P_{I,0}(t) = \sigma_0 \delta(t), \tag{44}$$

where $\delta(t)$ is the delta function, σ_0 is the initial probability vector that the system is in the state set $\{(I, 0, i) : 1 \leq i \leq m\}$ at time 0,

$$P_{I,k}(0) = 0, \quad k \geq 1, \tag{45}$$

$$P_{W,k}(0, x) = 0, \quad k \geq 0. \tag{46}$$

Let

$$\tilde{P}_{I,k}(s) = \int_0^{+\infty} e^{-st} P_{I,k}(t)\, dt, \;\; \tilde{P}_{W,k}(s, x) = \int_0^{+\infty} e^{-st} P_{W,k}(t, x)\, dt, \;\; k \geq 0,$$

and

$$\tilde{Q}_W^*(z, s, x) = \sum_{k=0}^{\infty} z^k \tilde{P}_{W,k}(s, x).$$

It follows from (41) and (42) that

$$\frac{\partial}{\partial x} \tilde{Q}_W^*(z, s, x) = \tilde{Q}_W^*(z, s, 0) \left\{ D^*(z) - [s + \alpha + \mu(x)] I \right\}.$$

Hence

$$\tilde{Q}_W^*(z, s, x) = \tilde{Q}_W^*(z, s, 0) \exp\left\{ D^*(z) x \right\} e^{-(s+\alpha)x} \bar{B}(x).$$

It is easy to see that

$$\tilde{P}_{W,k}(s, x) = \sum_{i=0}^{k} \tilde{P}_{W,i}(s, 0)\, P^D(k - i, x)\, e^{-(s+\alpha)x} \bar{B}(x), \;\; k \geq 0. \tag{47}$$

To determine the row vectors $\tilde{P}_{W,k}(s, 0)$ for $k \geq 0$, we introduce

$$\tilde{A}_k(s) = \int_0^{+\infty} P^D(k, x)\, e^{-(s+\alpha)x} dB(x), \;\; k \geq 0,$$

$$B_1^{(l)}(s) = \begin{pmatrix} D_0 - (s + l\theta) I & D_1 \\ \tilde{A}_0(s) & -I \end{pmatrix}, \;\; l \geq 0,$$

and

$$B_0^{(l)} = \begin{pmatrix} 0 & l\theta I \\ 0 & 0 \end{pmatrix}, \quad B_k(s) = \begin{pmatrix} 0 & D_k \\ \tilde{A}_{k-1}(s) & 0 \end{pmatrix}, \;\; k \geq 2.$$

We write

$$P_{IW}(s) = \left(\tilde{P}_{I,0}(s), \tilde{P}_{W,0}(s,0), \tilde{P}_{I,1}(s), \tilde{P}_{W,1}(s,0), \tilde{P}_{I,2}(s), \tilde{P}_{W,2}(s,0), \ldots\right).$$

It follows from (40), (43) to (47) that

$$P_{IW}(s)\, Q(s) = -(\hat{\sigma}, 0, 0, \ldots), \tag{48}$$

where $\hat{\sigma} = (\sigma_0, 0)$ and

$$Q(s) = \begin{pmatrix} B_1^{(0)}(s) & B_2(s) & B_3(s) & B_4(s) & \cdots \\ B_0^{(1)} & B_1^{(1)}(s) & B_2(s) & B_3(s) & \cdots \\ & B_0^{(2)} & B_1^{(2)}(s) & B_2(s) & \cdots \\ & & B_0^{(3)} & B_1^{(3)}(s) & \cdots \\ & & & \ddots & \ddots \end{pmatrix}. \tag{49}$$

To solve the equation (48), we write

$$Q_k(s) = \begin{pmatrix} B_1^{(k)}(s) & B_2(s) & B_3(s) & \cdots \\ B_0^{(k+1)}(s) & B_1^{(k+1)}(s) & B_2(s) & \cdots \\ & B_0^{(k+2)}(s) & B_1^{(k+2)}(s) & \cdots \\ & & \ddots & \ddots \end{pmatrix}, \quad k \geq 1.$$

We denote by $(\hat{Q}_{1,1}^{(k)}(s)^T, \hat{Q}_{2,1}^{(k)}(s)^T, \ldots)^T$ the first block-column of the fundamental matrix $\hat{Q}_k(s) = -Q_k^{-1}(s)$, which is the minimal nonnegative inverse of $-Q_k(s)$.

We define

$$R_j^{(k)}(s) = \sum_{i=1}^{\infty} B_{i+j}^{(k)}(s)\, \hat{Q}_{i,1}^{(k+1)}(s), \quad k \geq 0, j \geq 1, \tag{50}$$

$$G^{(k)}(s) = \hat{Q}_{1,1}^{(k)}(s)\, B_0^{(k)}(s), \quad k \geq 1, \tag{51}$$

and

$$U_k(s) = B_1^{(k)}(s) + R_1^{(k)}(s)\, B_0^{(k+1)}(s), \quad k \geq 0. \tag{52}$$

Similar discussions to that of Theorem 2, the following theorem is obvious.

Theorem 6.

$$Q(s) = [I - R_U(s)]\, U_D(s)\, [I - G_L(s)], \tag{53}$$

where

$$I - R_U(s) = \begin{pmatrix} I & -R_1^{(0)}(s) & -R_2^{(0)}(s) & -R_3^{(0)}(s) & \cdots \\ & I & -R_1^{(1)}(s) & -R_2^{(1)}(s) & \cdots \\ & & I & -R_1^{(2)}(s) & \cdots \\ & & & I & \cdots \\ & & & & \ddots \end{pmatrix},$$

$$U_D(s) = diag\,(U_0(s), U_1(s), U_2(s), U_3(s), \ldots),$$

$$I - G_L(s) = \begin{pmatrix} I & & & & \\ -G^{(1)}(s) & I & & & \\ & -G^{(2)}(s) & I & & \\ & & -G^{(3)}(s) & I & \\ & & & \ddots & \ddots \end{pmatrix}.$$

The equation (53) is called the RG-factorization of level-dependent Markov renewal processes of $M/G/1$ type. In fact, the RG-factorization holds for more general irreducible Markov renewal processes. Readers may refer to Li and Zhao (2004) for more details.

Let

$$X_1^{(l)}(s) = R_1^{(l)}(s), \; l \geq 0, \tag{54}$$

$$\begin{aligned} X_{k+1}^{(l)}(s) = {} & R_1^{(l)}(s)\,X_k^{(l+1)}(s) + R_2^{(l)}(s)\,X_{k-1}^{(l+2)}(s) \\ & + \cdots + R_k^{(l)}(s)\,X_1^{(l+k)}(s), \quad l \geq 0, \quad k \geq 1, \end{aligned} \tag{55}$$

and

$$Y_k^{(l)}(s) = G^{(l)}(s)\,G^{(l-1)}(s) \cdots G^{(l-k+1)}(s), \quad l \geq 1, \quad k \geq 1. \tag{56}$$

Lemma 5. If $s > 0$, then $I - R_U(s)$, $U_D(s)$ and $I - G_L(s)$ are invertible,

$$[I - R_U(s)]^{-1} = \begin{pmatrix} I & X_1^{(0)}(s) & X_2^{(0)}(s) & X_3^{(0)}(s) & \cdots \\ & I & X_1^{(1)}(s) & X_2^{(1)}(s) & \cdots \\ & & I & X_1^{(2)}(s) & \cdots \\ & & & I & \cdots \\ & & & & \ddots \end{pmatrix}, \tag{57}$$

$$U_D(s)^{-1} = \text{diag}\left(U_0(s)^{-1}, U_1(s)^{-1}, U_2(s)^{-1}, U_3(s)^{-1}, \ldots\right) \tag{58}$$

$$[I - G_L(s)]^{-1} = \begin{pmatrix} I & & & & \\ Y_1^{(1)}(s) & I & & & \\ Y_2^{(2)}(s) & Y_1^{(2)}(s) & I & & \\ Y_3^{(3)}(s) & Y_2^{(3)}(s) & Y_1^{(3)}(s) & I & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{59}$$

Proof: We first prove that the three matrices $I - R_U(s)$, $U_D(s)$ and $I - G_L(s)$ are invertible. Note that the two matrices $I - R_U(s)$ and $I - G_L(s)$ are obviously invertible, we only need to show that the matrix $U_D(s)$ is invertible. Since the Markov chain $Q(s)$ is transient for $s > 0$, the censored Markov chain $U_0(s)$ to level 0 is also transient. Hence, the matrix $U_0(s)$ is invertible. Based on the censoring property that the matrix $U_k(s)$ is invertible for each $k \geq 1$, it is easy to see that the matrix $U_D(s)$ is invertible.

Then, we provide the expressions for the inverses of the matrices $I - R_U(s)$, $U_D(s)$ and $I - G_L(s)$. To do this, we only prove (57) and the other two can be proved similarly. Noting that $[I - R_U(s)][I - R_U(s)]^{-1} = I$, by induction we can obtain

(i) the $(l, l+1)$st block-entry of $[I - R_U(s)]^{-1}$ is

$$X_1^{(l)}(s) = R_1^{(l)}(s), \; l \geq 0;$$

(ii) the $(l, l+k+1)$st block-entry of $[I - R_U(s)]^{-1}$ is

$$\begin{aligned} X_{k+1}^{(l)}(s) = {} & R_1^{(l)}(s) X_k^{(l+1)}(s) \\ & + R_2^{(l)}(s) X_{k-1}^{(l+2)}(s) + \cdots + R_k^{(l)}(s) X_1^{(l+k)}(s), \quad l \geq 0, k \geq 1. \end{aligned}$$

This completes the proof. □

Now, we solve equation $P_{IW}(s) Q(s) = -(\hat{\sigma}, 0, 0, \ldots)$. It follows from (53) and Lemma 5 that

$$\begin{aligned} P_{IW}(s) &= (\hat{\sigma}, 0, 0, \ldots)[I - G_L(s)]^{-1}[-U_D(s)]^{-1}[I - R_U(s)]^{-1} \\ &= \hat{\sigma}[-U_0(s)]^{-1}\left(I, X_1^{(0)}(s), X_2^{(0)}(s), \ldots\right). \end{aligned}$$

Therefore, we can express the vectors $\tilde{P}_{I,k}(s)$ and $\tilde{P}_{W,k}(s, 0)$ for all $k \geq 0$ as

$$\left(\tilde{P}_{I,0}(s), \tilde{P}_{W,0}(s, 0)\right) = \hat{\sigma}[-U_0(s)]^{-1}$$

and

$$\left(\tilde{P}_{I,k}(s), \tilde{P}_{W,k}(s, 0)\right) = \hat{\sigma}[-U_0(s)]^{-1} X_k^{(0)}(s), \quad k \geq 1.$$

Theorem 7. *The Laplace transform of the reliability function of the server is given by*

$$\tilde{\Re}(s) = \sum_{k=0}^{\infty} \left[\tilde{P}_{I,k}(s)\, e + \sum_{i=0}^{k} \tilde{P}_{W,i}(s,0)\, \tilde{C}_{k-i}(s)\, e \right],$$

where

$$\tilde{C}_k(s) = \int_0^{+\infty} e^{-(s+\alpha)} P^D(k,x)\, \bar{B}(x)\, dx.$$

The mean of the first failure time of the server is given by

$$MTTFF = \sum_{k=0}^{\infty} \left[\tilde{P}_{I,k}(0)\, e + \sum_{i=0}^{\infty} \tilde{P}_{W,i}(0,0)\, \tilde{C}_{k-i}(0)\, e \right]. \tag{60}$$

Proof: Noting that

$$\tilde{\Re}(s) = \sum_{k=0}^{\infty} \left[\tilde{P}_{I,k}(s)\, e + \int_0^{+\infty} \tilde{P}_{W,i}(s,x)\, dx e \right]$$

and

$$MTTFF = \tilde{\Re}(0),$$

simple computations lead to the proof. □

6. Queueing indices

In this section, we express the distributions of the stationary queue length and the busy period.

If the system is stable, we write

$$p_k = \lim_{t \to +\infty} P\{N(t) = k\}, \quad k \geq 0.$$

Theorem 8. *If the system is stable, then*

$$p_0 = \gamma \mathbf{x}_{0,1} e,$$

$$p_k = \gamma \left(\mathbf{x}_{k,1} e + \sum_{i=0}^{k-1} \mathbf{x}_{i,2} F_{k-1-i} + \alpha \sum_{j=0}^{k-1} \sum_{i=0}^{j} \mathbf{x}_{i,2} F_{j-i} H_{k-1-j} \right) e, \quad k \geq 1,$$

where F_k and H_k are defined in (23), γ and $\mathbf{x}_k = (\mathbf{x}_{k,1}, \mathbf{x}_{k,2})$ for $k \geq 0$ are given in (22) and Corollary 3, respectively.

Proof: If the system is stable, we write

$$p_k^{(I)} = \lim_{t\to+\infty} P\{L(t) = I, N(t) = k\}, \quad k \geq 0,$$

$$p_k^{(W)} = \lim_{t\to+\infty} P\{L(t) = W, N(t) = k\}, \quad k \geq 1,$$

$$p_k^{(R)} = \lim_{t\to+\infty} P\{L(t) = R, N(t) = k\}, \quad k \geq 1.$$

Noting that

$$p_0 = p_0^{(I)}, \quad p_k = p_k^{(I)} + p_k^{(W)} + p_k^{(R)}, \quad k \geq 1,$$

it follows from Theorem 4 that

$$p_0^{(I)} = p_{I,0}e, \; p_k^{(I)} = p_{I,k}e, \; k \geq 1,$$
$$p_k^{(W)} = \int_0^{+\infty} P_{W,k-1}(x)\, dxe, \; p_k^{(R)} = \int_0^{+\infty}\int_0^{+\infty} P_{R,k-1}(x, y)\, dxdy, k \geq 1.$$

Therefore, simple computations can complete this proof. □

Remark 8. The approach, provided in the above, for obtaining the distribution of the stationary queue length at an arbitrary epoch is different from that in Dudin and Klimenok (2000), where the distribution of the stationary queue length at the service completion epochs is first determined in terms of the embedded Markov chain, and the distribution of the stationary queue length at an arbitrary epoch is then determined by the associated Markov renewal process.

In what follows we consider the busy period. A busy period is defined as the period that starts at a time epoch at which a customer enters an empty system (where the server is free and there is no retrial customer) and ends at the departure epoch at which the system is left empty for first time. The busy period defined in this way consists of alternating periods of the generalized service period and the period during which the server is free and there exists at least one retrial customer in the orbit.

Let **B** be the length of a busy period, $\mathcal{B}(t) = P\{\mathbf{B} > t\}$ and $\tilde{\mathcal{B}}(s) = \int_0^{+\infty} e^{-st}\mathcal{B}(t)\, dt$. To derive the complementary distribution $\mathcal{B}(t)$, we first treat all the states $(I, 0, i)$ for $1 \leq i \leq m$ as an absorbing state '$*$', hence the busy period is a first passage time of the Markov process $\{(L(t), N(t), J(t), S(t), R(t)): t \geq 0\}$ defined in Section 2. Based on this, the computational procedure of $\tilde{\mathcal{B}}(s)$ is the same as that used in Theorem 7 for calculating the Laplace transform of the reliability function. For simplicity, we only give the main results with respect to $\tilde{\mathcal{B}}(s)$ while the detailed derivations are omitted.

Let

$$\tilde{A}_k^{\Psi}(s) = \int_0^{+\infty} P^{\Psi}(k, x)\, e^{-sx} dB(x), \; k \geq 0,$$

$$\mathbf{B}_1^{(k)}(s) = \begin{pmatrix} -I & \tilde{A}_1^{\Psi}(s) \\ (k+1)\theta & D_0 - [s + (k+1)\theta] I \end{pmatrix}, \; k \geq 0,$$

$$\mathbf{B}_0(s) = \begin{pmatrix} 0 & \tilde{A}_0^{\Psi}(s) \\ 0 & 0 \end{pmatrix}, \ \mathbf{B}_k(s) = \begin{pmatrix} 0 & \tilde{A}_k^{\Psi}(s) \\ D_{k-1} & 0 \end{pmatrix}, \ k \geq 2,$$

and

$$\mathbf{Q}(s) = \begin{pmatrix} \mathbf{B}_1^{(0)}(s) & \mathbf{B}_2(s) & \mathbf{B}_3(s) & \mathbf{B}_4(s) & \cdots \\ \mathbf{B}_0(s) & \mathbf{B}_1^{(1)}(s) & \mathbf{B}_2(s) & \mathbf{B}_3(s) & \cdots \\ & \mathbf{B}_0(s) & \mathbf{B}_1^{(2)}(s) & \mathbf{B}_2(s) & \cdots \\ & & \mathbf{B}_0(s) & \mathbf{B}_1^{(3)}(s) & \cdots \\ & & & \ddots & \ddots \end{pmatrix}.$$

For the matrix $\mathbf{Q}(s)$, we define the R-measure $\{\mathbf{R}_j^{(k)}\}$, the G-measure $\{\mathbf{G}^{(k)}\}$ and the matrix sequence $\{\mathbf{U}_k\}$, as defined in (50), (51) and (52), respectively. We write

$$\mathbf{X}_1^{(l)}(s) = \mathbf{R}_1^{(l)}(s), \ l \geq 0,$$

$$\mathbf{X}_{k+1}^{(l)}(s) = \mathbf{R}_1^{(l)}(s)\mathbf{X}_k^{(l+1)}(s) + \mathbf{R}_2^{(l)}(s)\mathbf{X}_{k-1}^{(l+2)}(s) + \cdots + \mathbf{R}_k^{(l)}(s)\mathbf{X}_1^{(l+k)}(s), \ l \geq 0, k \geq 1,$$

and

$$\mathbf{Y}_k^{(l)}(s) = \mathbf{G}^{(l)}(s)\mathbf{G}^{(l-1)}(s)\cdots\mathbf{G}^{(l-k+1)}(s), \quad l \geq 1, k \geq 1.$$

Let

$$\theta_B = (\theta_{B1}, \theta_{B2}, \theta_{B3}, \ldots),$$

with $\theta_{Bk} = (\frac{P_{I,0}D_k}{P_{I,0}D_{+}e}, 0)$ for $k \geq 1$ are of size $2m$,

$$\left(\tilde{P}_{W,0}(s,0), \tilde{P}_{I,1}(s)\right) = \left[\theta_{B1} + \sum_{i=2}^{\infty} \theta_{Bi} \mathbf{Y}_{i-1}^{(i-1)}(s)\right] [-\mathbf{U}_0(s)]^{-1}$$

and for $k \geq 2$,

$$(\tilde{P}_{W,k-1}(s,0), \tilde{P}_{I,k}(s)) = \sum_{l=2}^{k-1} \left[\theta_{Bl} + \sum_{i=l+1}^{\infty} \theta_{Bi+l} \mathbf{Y}_{i-l}^{(i-1)}(s)\right] [-\mathbf{U}_{l-1}(s)]^{-1} \mathbf{X}_{k-l}^{(l-1)}(s)$$
$$+ \left[\theta_{Bk} + \sum_{i=k+1}^{\infty} \theta_{Bi} \mathbf{Y}_{i-k}^{(i-k)}(s)\right] [-\mathbf{U}_{k-1}(s)]^{-1}.$$

Theorem 9. *The Laplace transform of $\mathcal{B}(t)$ is given by*

$$\tilde{\mathcal{B}}(s) = \sum_{k=1}^{\infty} \tilde{P}_{I,k}(s)\, e + \sum_{k=0}^{\infty} \sum_{i=0}^{k} \tilde{P}_{W,i}(s,0)\, \tilde{C}_{k-i}(s)\, e$$

$$+ \sum_{l=0}^{\infty} \sum_{k=0}^{l} \sum_{i=0}^{k} \tilde{P}_{W,i}(s,0)\, \tilde{C}_{k-i}(s)\, \tilde{D}_{l-k}(s)\, e,$$

where

$$\tilde{C}_k(s) = \int_0^{+\infty} e^{-sx} P^{\Psi}(k,x)\, \bar{B}(x)\, dx, \ \ \tilde{D}_k(s) = \int_0^{+\infty} e^{-sx} P^{D}(k,y)\, \bar{V}(y)\, dy.$$

The mean of the busy period is given by

$$E[\mathbf{B}] = \sum_{k=1}^{\infty} \tilde{P}_{I,k}(0)\, e + \sum_{k=0}^{\infty} \sum_{i=0}^{k} \tilde{P}_{W,i}(0,0)\, \tilde{C}_{k-i}(0)\, e$$

$$+ \sum_{l=0}^{\infty} \sum_{k=0}^{l} \sum_{i=0}^{k} \tilde{P}_{W,i}(0,0)\, \tilde{C}_{k-i}(0)\, \tilde{D}_{l-k}(0)\, e. \tag{61}$$

Remark 9. Theorem 9 applies the RG-factorization of a level-dependent Markov renewal process of $M/G/1$ type to provide a novel method for calculating the Laplace transform of the complementary distribution of the busy period in a block-structured queue, although this method is standard for an $M/G/1$ (retrial) queue.

7. Two algorithms

In this section, we provide two algorithms: The first one is used to compute the stationary probability vector of the continuous-time level-dependent Markov chain of $M/G/1$ type, and the second one is to calculate the mean of the first passage time of this Markov chain. The two algorithms are based on the R-measure, which is a key to express both the stationary probability vector and the Laplace transform of a first passage time.

In the class of asymptotically quasi-Toeplitz 2-dimensional Markov chains (see Section 4 in Dudin and Klimenok, 2000), Neuts and Rao (1990) proposed an approximate algorithm to compute the stationary probability vector of a positive recurrent level-dependent QBD process. With this algorithm, one first needs to modify the level-dependent QBD process to a corresponding level-independent QBD process; then use the matrix-geometric solution of the modified process to approximate the stationary probability vector. Dudin and Klimenok (2000) generalized this algorithm to deal with a level-dependent Markov chain of $M/G/1$ type. When the level-dependent Markov chain is modified as the level-independent case, they used the censoring technique to construct a system of finitely-many linear equations whose solution approximates the stationary probability vector.

Our algorithm is different from that of Dudin and Klimenok (2000), although both algorithms use the same censoring technique. The algorithm given in Dudin and Klimenok (2000) is a direct truncation method. One drawback of the direct truncation method is that computational accuracy heavily depends on the truncation size of the censored matrix for some heavy-traffic cases. Noting that in this paper the stationary probability vector of a positive recurrent Markov chain of $M/G/1$ type can be expressed by the R-measure, our algorithm effectively use all the subvectors in the stationary probability vector, including those corresponding to the censored chain as in Dudin and Klimenok (2000). Therefore, our algorithm improves that in Dudin and Klimenok (2000). For a probabilistic interpretation of the improved algorithm, we would like to note that Neuts and Rao (1990) proposed a similar way for improving the direct truncation method by using the matrix-geometric solution.

For a level-dependent QBD process, which may not be a asymptotically quasi-Toeplitz 2-dimensional Markov chain, Bright and Taylor (1995) proposed a method based on sample path truncations to compute the truncated R-measure $\{R_k, 1 \leq k \leq K^* - 1\}$ and the truncated stationary probability vector $\{\pi_k, 1 \leq k \leq K^*\}$. Also, they illustrated how to choose the number K^*. Bright and Taylor (1995) improved the direct truncation method.

We now present the first algorithm in five steps:

Step 1. A crucial number

We need to modify matrix Q in (20) to a level-independent infinitesimal generator. Let N be a positive integer such that the modified Markov chain beginning from level N is level-independent. The following rule determines the number N.

Let G_n be the minimal nonnegative solution to the matrix equation

$$A_0^{(n)} + A_1^{(n)} G_n + \sum_{k=2}^{\infty} A_k G_n^k = 0. \tag{62}$$

Using (62) and letting $n \to \infty$, we obtain

$$\begin{pmatrix} \frac{1}{n\theta} I & \\ & I \end{pmatrix} A_0^{(n)} \to \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{n\theta} I & \\ & I \end{pmatrix} A_1^{(n)} \to \begin{pmatrix} -I & 0 \\ C_0 & -I \end{pmatrix},$$

$$\begin{pmatrix} \frac{1}{n\theta} I & \\ & I \end{pmatrix} A_k \to \begin{pmatrix} 0 & 0 \\ C_{l-1} & 0 \end{pmatrix}, \ k \geq 2.$$

Let G_∞ be the minimal nonnegative solution to the equation

$$\begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} -I & 0 \\ C_0 & -I \end{pmatrix} G_\infty + \sum_{l=2}^{\infty} \begin{pmatrix} 0 & 0 \\ C_{l-1} & 0 \end{pmatrix} (G_\infty)^l = 0. \tag{63}$$

Then $G_\infty = \lim_{n\to\infty} G_n$. Hence, there always exists a sufficiently large positive integer N such that for all $n \geq N$, $||G_n - G_\infty|| < \varepsilon$, where $\varepsilon > 0$ is a predetermined smaller number and $|| \cdot ||$ denotes norm of matrix.

Step 2. A modified level-independent Markov chain of *M/G/1* type

Once N is found in Step one, the matrix Q can be modified to

$$\tilde{Q}_N = \begin{pmatrix} \mathbf{C}_1^{(N-1)} & \mathbf{C}_2^{(N-1)} & \mathbf{C}_3^{(N-1)} & \mathbf{C}_4^{(N-1)} & \cdots \\ \mathbf{C}_0^{(N)} & A_1^{(N)} & A_2 & A_3 & \cdots \\ & A_0^{(N)} & A_1^{(N)} & A_2 & \cdots \\ & & A_0^{(N)} & A_1^{(N)} & \cdots \\ & & & A_0^{(N)} & \cdots \\ & & & & \ddots \end{pmatrix}, \tag{64}$$

where

$$\mathbf{C}_0^{(N)} = \left(0, 0, \ldots, 0, A_0^{(N)}\right),$$

$$\mathbf{C}_1^{(N-1)} = \begin{pmatrix} A_1^{(0)} & A_2 & \cdots & A_{N-1} & A_N \\ A_0^{(1)} & A_1^{(1)} & \cdots & A_{N-2} & A_{N-1} \\ & A_0^{(2)} & \cdots & A_{N-3} & A_{N-2} \\ & & \ddots & \vdots & \vdots \\ & & & A_0^{(N-1)} & A_1^{(N-1)} \end{pmatrix}, \quad \mathbf{C}_k^{(N-1)} = \begin{pmatrix} A_{N+k-1} \\ A_{N+k-2} \\ A_{N+k-3} \\ \vdots \\ A_k \end{pmatrix}, k \geq 2.$$

Step 3. The R-measure

Using Theorem 1 in Li and Zhao (2002) we define

$$U = A_1^{(N)} + \sum_{k=1}^{\infty} A_{k+1} G_\infty^k$$

By (18) and (19) in Li and Zhao (2002), we have the R-measure

$$R_{0,k} = \sum_{i=1}^{\infty} \mathbf{C}_{k+i}^{(N-1)} G_\infty^{i-1} (-U)^{-1}$$

and

$$R_k = \sum_{i=1}^{\infty} A_{k+i} G_\infty^{i-1} (-U)^{-1}, \quad k \geq 1.$$

Step 4. The matrix U_0 and the vector $\mathbf{x}_0$

According to Theorem 1 in Li and Zhao (2002) we define

$$U_0 = \mathbf{C}_1^{(N-1)} + \sum_{k=1}^{\infty} \mathbf{C}_{k+1}^{(N-1)} G_\infty^{k-1} (-U)^{-1} \mathbf{C}_0^{(N)}.$$

It is clear that U_0 is the infinitesimal generator of the positive recurrent censored Markov chain. Let $\mathbf{x}_0$ be the stationary probability vector of this censored Markov chain to level 0. Then $\mathbf{x}_0 U_0 = 0$ and $\mathbf{x}_0 e = 1$.

Let $\mathbf{W} = (U_0 : e)$. Then $\mathbf{x}_0 \mathbf{W} = (0 : 1)$ according to $\mathbf{x}_0 U_0 = 0$ and $\mathbf{x}_0 e = 1$. Therefore, we obtain

$$\mathbf{x}_0 = (0 : 1)\, \mathbf{W}^T \left(\mathbf{W}\mathbf{W}^T\right)^{-1}.$$

We may also use a stable algorithm to calculate the vector $\mathbf{x}_0$, which is an iterative method based on sparse matrix algorithms. Let $U_0 = W - V$, where the matrix W is needed to be invertible. Thus, we obtain that $\mathbf{x}_0 W = \mathbf{x}_0 V$. This suggests the following iterative scheme for computing $\mathbf{x}_0$.

$$\mathbf{x}_0\,(0) = (\alpha_1, \alpha_2, \ldots, \alpha_{Nm})$$

and

$$\mathbf{x}_0\,(r+1)\, W = \mathbf{x}_0\,(r)\, V,$$

where $\alpha_k \geq 0$ and $\sum_{k=1}^{Nm} \alpha_k = 1$. It is clear that $\mathbf{x}_0 = \lim_{r\to\infty} \mathbf{x}_0\,(r)$.

Step 5. The stationary probability vector

We denote by $(\tilde{\pi}_0, \tilde{\pi}_1, \tilde{\pi}_2, \ldots)$ the stationary probability vector of the modified Markov chain $\tilde{Q}_N$, where the size of $\tilde{\pi}_0$ is Nm and the size of $\tilde{\pi}_k$ is m for $k \geq 1$. Using Subsection 5.1 of Li and Zhao (2002), we obtain

$$\tilde{\pi}_0 = c\, \mathbf{x}_0,$$

$$\tilde{\pi}_1 = \tilde{\pi}_0 R_{0,1},$$

$$\tilde{\pi}_k = \tilde{\pi}_0 R_{0,k} + \sum_{i=1}^{k-1} \tilde{\pi}_i R_{k-i},\ k \geq 2,$$

where

$$c = \frac{1}{1 + \mathbf{x}_0 R_0 (I - R)^{-1} e},$$

with

$$R_0 = \sum_{k=1}^{\infty} R_{0,k} \ \text{ and } \ R = \sum_{k=1}^{\infty} R_k.$$

Remark 10. Since the matrix G_∞ can be numerically calculated from the matrix equation (63) in terms of the well-known algorithms, e.g., see Neuts (1989, 1995), the matrices U_0, $R_{0,k}$ and R_k for $k \geq 1$ can be numerically given. As such, the stationary probability vector can be provided. Therefore, the algorithm based on the R-measure is different from that in Dudin and Klimenok (2000).

In what follows we estimate the computational complexity of the above algorithm. We use Horner's algorithm for matrix polynomials, and then measure the complexity by the total number of multiplications and divisions of two floating point numbers. It is readily seen that the computational complexity of the matrix G_n (or G_∞) of size m is $O(m^3)$. If we assume that the inverse of the matrix of size m is performed by some version of Gaussian elimination, then it is well-known that the complexity of computing this inverse is $O(m^3)$. Furthermore, we can easily give the complexity of computing the matrices U, R_k, $R_{0,k}$, U_0 and the vectors x_0 and π_k. Let N be the truncated number given in Step one of this algorithm. Note that the size of the matrix U_0 is Nm, it is clear from Step five in this algorithm that the complexity of computing the vectors π_k for $k \geq 1$ is $O(N^3m^3)$. It is easy to check that the computational complexity of the algorithm presented in Dudin and Klimenok (2000) is $O(\mathbf{N}^3m^3)$, where the number $\mathbf{N}$ is the size of the censored matrix (see (12) in Dudin and Klimenok, 2000). Note that N is not more than $\mathbf{N}$ under the same precision, thus our algorithm improves the computational complexity of the algorithm given in Dudin and Klimenok (2000).

We now describe the second algorithm for calculating the mean of the first passage time of the level-dependent Markov chain of $M/G/1$ type. This algorithm can be modified to compute high-order moments of the first passage time. For simplicity of description, we only discuss the first failure time of the server.

It is seen from (60) that the mean of the first failure time of the server only depends on the matrix $Q(s)$ of (49) at $s = 0$. Hence, we modify the level-dependent Markov chain $Q(0)$ of $M/G/1$ type to its corresponding level-independent Markov chain, where the positive integer N can be chosen by the same method as that using (62) and (63). Furthermore, the matrix $Q(0)$ is modified as

$$\mathbf{Q}_N(0) = \begin{pmatrix} \mathbf{D}_1^{(N-1)}(0) & \mathbf{D}_2^{(N-1)}(0) & \mathbf{D}_3^{(N-1)}(0) & \mathbf{D}_4^{(N-1)}(0) & \cdots \\ \mathbf{D}_0^{(N)}(0) & B_1^{(N)}(0) & B_2(0) & B_3(0) & \cdots \\ & B_0^{(N)}(0) & B_1^{(N)}(0) & B_2(0) & \cdots \\ & & B_0^{(N)}(0) & B_1^{(N)}(0) & \cdots \\ & & & B_0^{(N)}(0) & \cdots \\ & & & & \ddots \end{pmatrix}, \tag{65}$$

where

$$\mathbf{D}_0^{(N)}(0) = \left(0, 0, \ldots, 0, B_0^{(N)}(0)\right),$$

$$\mathbf{D}_1^{(N-1)} = \begin{pmatrix} B_1^{(0)}(0) & B_2(0) & \cdots & B_{N-1}(0) & B_N(0) \\ B_0^{(1)}(0) & B_1^{(1)}(0) & \cdots & B_{N-2}(0) & B_{N-1}(0) \\ & B_0^{(2)}(0) & \cdots & B_{N-3}(0) & B_{N-2}(0) \\ & & \ddots & \vdots & \vdots \\ & & & B_0^{(N-1)}(0) & B_1^{(N-1)}(0) \end{pmatrix}$$

and

$$\mathbf{D}_k^{(N-1)}(0) = \begin{pmatrix} B_{N+k-1}(0) \\ B_{N+k-2}(0) \\ B_{N+k-3}(0) \\ \vdots \\ B_k(0) \end{pmatrix}, \quad k \geq 2.$$

Let $\mathbf{G}$ be the minimal nonnegative solution to the matrix equation

$$B_0^{(N)}(0) + B_1^{(N)}(0)\,\mathbf{G} + \sum_{k=2}^{\infty} B_k(0)\,\mathbf{G}^k = 0.$$

Using Theorem 1 in Li and Zhao (2002), we define

$$\mathbf{U} = B_1^{(N)}(0) + \sum_{k=1}^{\infty} B_{k+1}(0)\,\mathbf{G}^k$$

and

$$\mathbf{U}_0 = \mathbf{D}_1^{(N-1)}(0) + \sum_{k=1}^{\infty} \mathbf{D}_{k+1}^{(N-1)}(0)\,\mathbf{G}^{k-1}(-\mathbf{U})^{-1}\,\mathbf{D}_0^{(N)}(0).$$

By (18) and (19) in Li and Zhao (2002), we define a R-measure as

$$\mathbf{R}_{0,k} = \sum_{i=1}^{\infty} \mathbf{D}_{k+i}^{(N-1)}(0)\,\mathbf{G}^{i-1}(-\mathbf{U})^{-1}, \quad k \geq 1,$$

and

$$\mathbf{R}_k = \sum_{i=1}^{\infty} B_{k+i}(0)\,\mathbf{G}^{i-1}(-\mathbf{U})^{-1}, \quad k \geq 1.$$

Let

$$X_1^{(0)} = \mathbf{R}_{0,1}, \quad X_1^{(l)} = \mathbf{R}_1, \; l \geq 1,$$

$$X_{k+1}^{(0)} = \mathbf{R}_{0,1} X_k^{(1)} + \mathbf{R}_{0,2} X_{k-1}^{(2)} + \cdots + \mathbf{R}_{0,k} X_1^{(k)}, \quad k \geq 1.$$

We denote by $\hat{\sigma}$ a probability vector of size m. Note that the size of the censored matrix $\mathbf{U}_0$ is Nm, we write

$$(\mathcal{P}_{I,0}, \mathcal{P}_{W,0}, \mathcal{P}_{I,1}, \mathcal{P}_{W,1}, \ldots, \mathcal{P}_{I,N-1}, \mathcal{P}_{W,N-1}) = \Bigg(\hat{\sigma}, \quad \underbrace{0, 0, \ldots, 0}_{N-1 \text{ vector } 0 \text{ of size } m}\Bigg)(-\mathbf{U}_0)^{-1},$$

$$\left(\mathcal{P}_{I,N+k-1}, \mathcal{P}_{W,N+k-1}\right) = \Bigg(\hat{\sigma}, \quad \underbrace{0, 0, \ldots, 0}_{N-1 \text{ vector } 0 \text{ of size } m}\Bigg)(-\mathbf{U}_0)^{-1}\, X_k^{(0)}, \quad k \geq 1,$$

and

$$\mathcal{C}_k = \int_0^{+\infty} e^{-\alpha} P^D(k, x)\, \bar{B}(x)\, dx.$$

Then the mean of the first failure time of the server is given by

$$MTTFF = \sum_{k=0}^{\infty} \left[\mathcal{P}_{I,k} e + \sum_{i=0}^{\infty} \mathcal{P}_{W,i} \mathcal{C}_{k-i} e \right].$$

Note that the computational complexity of the *MTTFF* is $O\left(N^3 m^3\right)$.

8. Numerical examples

In this section, we give some simple numerical examples to illustrate how the two algorithms work.

The most laborious part in implementing our approximate algorithm is to compute the matrix U_0 and the R-measure $\left\{R_{0,k}\right\}$ and $\{R_k\}$. According to their expressions, we can numerically obtain the stationary probability vector $\{\tilde{\pi}_0, \tilde{\pi}_1, \tilde{\pi}_2, \ldots\}$ of the modified Markov chain. Let $(\pi_0, \pi_1, \ldots, \pi_{N-1}) = \tilde{\pi}_0$ and $\pi_{N+k-1} = \tilde{\pi}_k$ for $k \geq 1$. Then the mean and variance of the stationary orbit size including that customer being served are approximately given by

$$L = \sum_{k=1}^{\infty} k \pi_k e \ \text{ and } \ V = \sum_{k=1}^{\infty} k^2 \pi_k e - L^2,$$

respectively.

For the repairable *BMAP/G/1* retrial queue, we let $m = 2$ and

$$D_0 = \begin{pmatrix} -(\lambda + 2) & \lambda \\ 1 & -5 \end{pmatrix}, \ D_1 = \begin{pmatrix} 2 & 0 \\ 2 & 2 \end{pmatrix}, \ D_k = 0, k \geq 2,$$

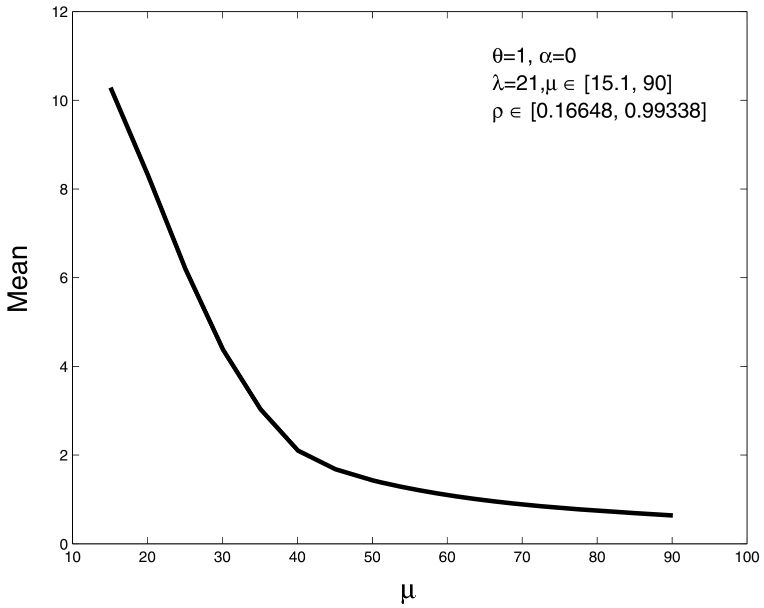

Fig. 1 The mean L depends on μ

where $\lambda > 0$. Obviously, the stationary arrival rate of the MAP is give by

$$r^* = 4 - \frac{6}{3 + \lambda} < 4.$$

Let

$$B(t) = \int_0^t \frac{\mu (\mu x)^3}{3!} e^{-\mu x} dx.$$

Then the service time distribution is Erlang with expression $E_4(\mu)$, and also the mean of the service time is $\frac{4}{\mu}$. To simplify our numerical analysis, we consider the case with $\alpha = 0$, which shows that the server is always good and reliable, and numerically discuss the mean and variance of the stationary queue length. It is not difficult that we may discuss other performance measures in the queue and even more complicated queueing systems. It is easy to check that the queue is stable if and only if $\rho = \frac{8(3+2\lambda)}{\mu(3+\lambda)} < 1$.

Now, we provide some numerical discussions for the queueing model. The first example is one-dimensional and the second one is two-dimensional.

Example 1. We use the following parameter values: $\theta = 1$, $\lambda = 21$ and $\mu \in [15.1, 90]$. In this case, it is easy to check $\rho \in [0.167, 0.993]$. Figures 1 and 2 illustrate the dependence of the mean L and the variance V on the parameter μ or the traffic intensity ρ, respectively. Figures 1 and 2 shows that the mean L and the variance V are all decreasing in μ.

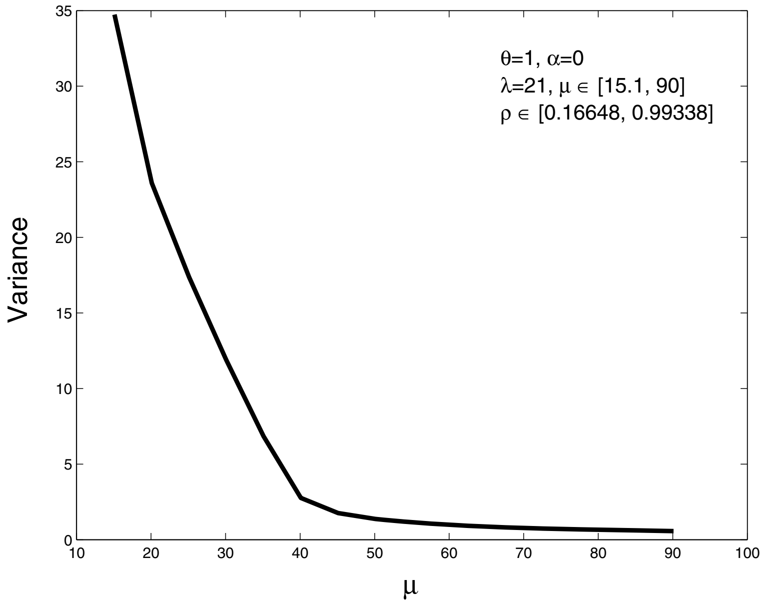

Fig. 2 The variance V depends on μ

Example 2. $\theta = 1$, $\lambda \in [1, 10]$ and $\mu \in [15, 90]$. In this case, it is easy to check $\rho \in [0.11, 0.94]$. Figure 3 shows that the mean L is decreasing in μ and increasing in λ. Figure 4 illustrates that the variance V is decreasing in μ and increasing in λ.

These examples validate our algorithms for computing the stationary probability vector of the Markov chain Q of $M/G/1$ type. Moreover, our algorithms can conveniently deal with the case of heavy traffic intensity. This is one reason why the introduction of the R-measure $\left\{R_{0,k}\right\}$ and $\{R_k\}$ to our algorithm improves the algorithm given in Dudin and Klimenok (2000). Also, in these examples, the truncated number $N = 20$ is sufficiently large to guarantee the computational precision.

9. Concluding remarks

In this paper, we considered the *BMAP/G/1* retrial queue with server breakdowns and repairs. We obtained (i) the reliability indices of the server: The stationary availability, the stationary failure frequency and the reliability function, and (ii) the queueing indices of the system: The stationary queue length and the busy period. In principle, the indices obtained in this paper can be numerically computed based on the matrix-analytic method extensively studied during the past two decades.

For retrial queueing systems of $M/G/1$ type, we showed that they can be studied by means of continuous-time level-dependent Markov chains of $M/G/1$ type and level-dependent Markov renewal processes of $M/G/1$ type. We provided the RG-factorizations for both continuous-time level-dependent Markov chains of $M/G/1$ type and for level-dependent Markov renewal processes of $M/G/1$ type. The RG-factorizations can be applied not only to deal with the stationary probability vector and

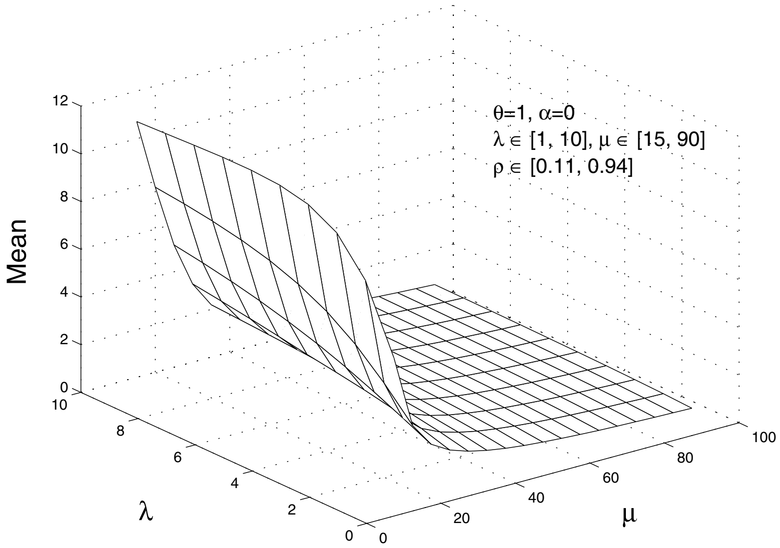

Fig. 3 The mean L depends on λ and μ

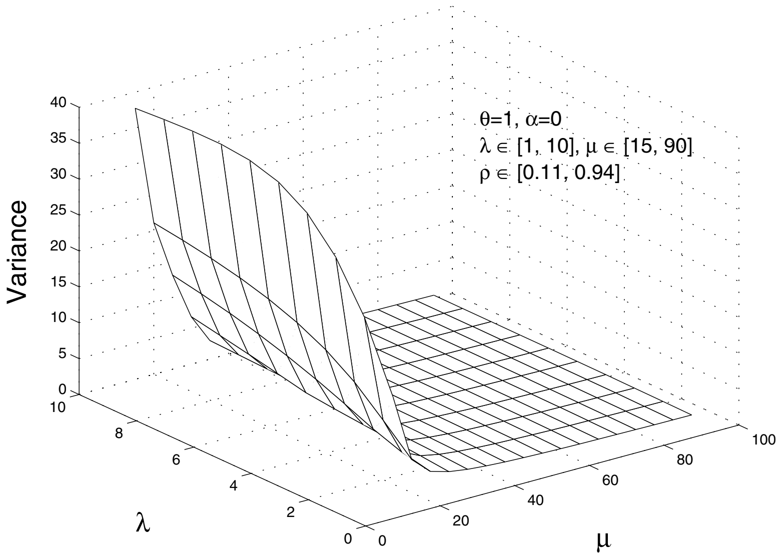

Fig. 4 The variance V depends on λ and μ

the associated indices such as the stationary availability, failure frequency and queue length, but also to discuss the first passage times and the associated indices such as the reliability function and the busy period. We expect that the approach developed in this paper can also be used to study other stochastic models with a level-dependent block structure.

Acknowledgments The authors thank the two referees for their valuable comments and remarks, and acknowledge that this work was supported by a research grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the National Natural Science Foundation of China (Grant No. 90412012). Dr. Li also acknowledges the support from Carleton University.

References

Aissani, A.A. (1994). "Retrial Queue with Redundancy and Unreliable Server." *Queueing Systems* 17, 431–449.

Aissani, A. and J.R. Artalejo. (1998). "On the Single Server Retrial Queue Subject to Breakdowns." *Queueing Systems* 30, 309–321.

Artalejo, J.R. (1994). "New Results in Retrial Queueing Systems with Breakdowns of the Servers." *Statistics Neerlandica* 48, 23–36.

Artalejo, J.R. (1999). "A Classified Bibliography of Research on Retrial Queues: Progress in 1990–1999." *Top* 7, 187–211.

Artalejo, J.R. and A. Gómez-Corral. (1998). "Unreliable Retrial Queues due to Service Interruptions Arising from Facsimile Networks." *Belg J Oper Res Statist Comput Sci* 38, 31–41.

Asmussen, S. (1987). *Applied Probability and Queues*. John Wiley & Sons.

Breuer, L., A. Dudin, and V. Klimenok. (2002). "A Retrial *BMAP/PH/N* System." *Queueing Systems* 40, 433–457.

Bright, L. and P.G. Taylor. (1995). "Calculating the Equilibrium Distribution in Level Dependent Quasi-birth-and-death Processes." *Stochastic Models* 11, 497–525.

Cao, J. and K. Chen. (1982). "Analysis of the $M/G/1$ Queueing System with Repairable Service Station." *Acta Math. Appl. Sinica* 5, 113–127.

Chakravarthy, S.R. (2001). "The Batch Markovian Arrival Processes: A Review and Future Work." In A. Krishnamoorthy et al. (eds.), *Advances in Probability Theory and Stochastic Processes*. Notable Publications, Inc., New Jersey, pp. 21–49.

Chakravarthy, S.R. and A. Dudin. (2002). "A Multi-server Retrial Queue with BMAP Arrivals and Group Services." *Queueing Systems* 42, 5–31.

Chakravarthy, S.R. and A. Dudin. (2003). "Analysis of a Retrial Queuing Model with MAP Arrivals and Two Types of Customers." *Math. Comput. Modelling* 37, 343–363.

Choi, B.D., Y. Yang, and B. Kim, (1999). "$MAP1/MAP2/M/c$ Retrial Queue with Guard Channels and its Applications to Cellular Networks." *Top* 7, 231–248.

Choi, B.D., Y.H. Chung, and A. Dudin. (2001). "The *BMAP/SM/1* Retrial Queue with Controllable Operation Modes." *European J. Oper. Res.* 131, 16–30.

Diamond, J.E. and A.S. Alfa. (1995). "Matrix Analytic Methods for $M/PH/1$ Retrial Queues with Buffers." *Stochastic Models* 11, 447–470.

Diamond, J.E. and A.S. Alfa (1998). "Matrix Analytic Methods for $MAP/PH/1$ Retrial Queues with Buffers." *Stochastic Models* 14, 1151–1187.

Diamond, J.E. and A.S. Alfa. (1999). "Matrix Analytic Methods for Multi-server Retrial Queues with Buffers." *Top* 7, 249–266.

Dudin, A. and V. Klimenok (1999). "*BMAP/SM/1* Model with Markov Modulated Retrials." *Top* 7, 267–278.

Dudin, A. and V. Klimenok. (2000). "A Retrial *BMAP/SM/1* System with Linear Requests." *Queueing Systems* 34, 47–66.

Falin, G.I. (1990). "A Survey of Retrial Queues." *Queueing Systems* 7, 127–167.

Falin, G.I. and J.G.C. Templeton. (1997). *Retrial Queues*. Chapman & Hall: London.

Ferng, H.W. and J.F. Chang. (2001). "Departure Processes of *BMAP/G/1* Queues." *Queueing Systems* 39, 109–135.

Gnedenko, B.V. and I.N. Kovalenko. (1989). *Introduction to Queueing Theory, Second Edition*. Birkhauser: Boston.

Grassmann, W.K. and D.P. Heyman. (1990). "Equilibrium Distribution of Block-structured Markov Chains with Repeating Rows." *J. Appl. Prob.* 27, 557–576.

He, Q., H. Li, and Y.Q. Zhao. (2000). "Ergodicity of the *BMAP/PH/s/s+k* Retrial Queue with PH-retrial Times." *Queueing Systems* 35, 323–347.

Horn, R.A and C.R. Johnson. (1985). *Matrix Analysis*. Cambridge University Press: London.

Hsu, G.H., Yuan, X. and Li, Q.L. (2000). "First Passage Times for Markov Renewal Processes and Application." *Science in China, Series A* 43, 1238–1249.

Latouche, G. and V. Ramaswami. (1999). *Introduction to Matrix Analytic Methods in Stochastic Modeling*. SIAM, Philadelphia.

Lee, G. and J. Jeon. (1999). "Analysis of an $N/G/1$ Finite Queue with the Supplementary Variable Method." *J Appl Math Stochastic Anal* 12, 429–434.

Lee, G. and Jeon, J. (2000). "A New Approach to an $N/G/1$ Queue." *Queueing Systems* 35, 317–322.

Li, Q.L. (1996). "Queue System $M/SM(PH/SM)/1$ with Repairable Service Station." *Mathematical Applicata* 9, 422–428.

Li, Q.L. and J. Cao. (2000). "The Repairable Queueing System $MAP/PH(M/PH)/2$ with Dependent Repairs." *Systems Sciences and Mathematical Science* 20, 78–86.

Li, Q.L., M. Tan, and Y. Sun. (1999). "The $SM/PH/1$ Queue with Repairable Server of PH Lifetime." In *IFAC 14th Triennial World Congress*, *Vol. A*. Beijing, P.R. China, 297–305.

Li, Q.L., D.J. Xu, and J. Cao. (1997). "Reliability Approximation of a Markov Queueing System with Server Breakdown and Repair." *Microelectron. Reliab.* 37, 1203–1212.

Li, Q.L. and Y.Q. Zhao. (2002). "A Constructive Method for Finding β-invariant Measures for Transition Matrices of $M/G/1$ type." In G. Latouche and P.G. Taylor (eds.), *Matrix Analytic Methods Theory and Applications*. World Scientific, pp. 237–264.

Li, Q.L. and Y.Q. Zhao. (2004). "The RG-factorization in block-structured Markov renewal process with applications." In X. Zhu (ed.), *Observation, Theory and Modeling of Atmospheric Variability*. World Scientific, pp. 545–568.

Li, Q.L. and Y.Q. Zhao. (2003). "β-invariant Measures for Transition Matrices of $GI/M/1$ type." *Stochastic Models* 19, 201–233.

Li, W., D.H. Shi, and X.L. Chao. (1997). "Reliability Analysis of $M/G/1$ Queueing Systems with Server Breakdowns and Vacations." *J. Appl. Prob.* 34, 546–555.

Liang, H.M. and V.G. Kulkarni. (1993). "Stability Condition for a Single Retrial Queue." *Adv. in Appl. Prob.* 25, 690–701.

Lucantoni, D.M. (1991). "New Results on the Single Server Queue with a Batch Markovian Arrival Process." *Stochastic Models* 7, 1–46.

Lucantoni, D.M. (1993). "The *BMAP/G/1* queue: A tutorial." In L. Donatiello and R. Nelson (eds.), *Models and Techniques for Performance Evaluation of Computer and Communication Systems* Springer-Verlag: New York.

Lucantoni, D.M., G.L. Choudhury, and W. Whitt. (1994). "The Transient *BMAP/G/1* Queue." *Stochastic Models* 10, 145–182.

Lucantoni, D.M. and M.F. Neuts. (1994). "Some Steady-state Distributions for the $MAP/SM/1$ Queue." *Stochastic Models* 10, 575–598.

Kulkarni, V.G. and B.D. Choi. (1990). "Retrial Queues with Server Subject to Breakdowns and Repairs." *Queueing Systems* 7, 191–208.

Kulkarni, V.G. and H.M. Liang. (1997). "Retrial Queues Revisited." In J.H. Dshalalow (ed.), *Frontiers in Queueing: Models and Applications in Science and Engineering*. CRC Press: Boca Raton, FL, pp. 19–34.

Mitrany, I.L. and B. Avi-Ttzhak. (1968). "A Many Server Queue with Service Interruptions." *Operations Research* 16, 628–638.

Neuts, M.F. (1989). *Structured Stochastic Matrices of $M/G/1$ Type and Their Applications*. Marcel Decker Inc.: New York.

Neuts, M.F. (1995). "Matrix-analytic methods in the Theory of queues." In J.H. Dshalalow (ed.), *Advances in Queueing: Theory, Methods and Open Problems*. CRC Press: Boca Raton, FL, pp. 265–292.

Neuts, M.F. and D.M. Lucantoni. (1979). "A Markovian Queue with N Servers Subject to Breakdowns and Repairs." *Managm. Sci.* 25, 849–861.

Neuts, M.F. and B.M. Rao. (1990). "Numerical Investigation of a Multiserver Retrial Model." *Queueing Systems* 7, 169–190.

Ramaswami, V. (1980). "The $N/G/1$ Queue and its Detailed Analysis." *Adv. Appl. Prob.* 12, 222–261.

Wang, J., J. Cao, and Q.L. Li. (2001). "Reliability Analysis of the Retrial Queue with Server Breakdowns and Repairs." *Queueing Systems* 38, 363–380.

Yang, T. and H. Li. (1994). "The $M/G/1$ Retrial Queue with the Server Subject to Starting Failures." *Queueing Systems* 16, 83–96.

Yang, T. and J.G.C. Templeton. (1987). "A Survey on Retrial Queues." *Queueing Systems* 2, 201–233.

Zhao, Y.Q. (2000). "Censoring technique in studying block-structured Markov chains." In G. Latouche and P.G. Taylor (eds.), *Advances in Algorithmic Methods for Stochastic Models*. Notable Publications Inc.: New Jersey, pp. 417–433.