J. Math. Anal. Appl. 436 (2016) 556–567

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

The M/M/C queueing system in a random environment

Zaiming Liu [1], Senlin Yu *,[1]

School of Mathematics and Statistics, Central South University, Changsha 410083, Hunan, PR China

ARTICLE INFO

Article history:
Received 30 June 2015
Available online 21 December 2015
Submitted by V. Pozdnyakov

Keywords:
Random environment
Markov modulated
M/M/C system
Applied probability

ABSTRACT

An $M/M/C$ queueing system operating in a Markovian environment is studied. This paper focuses on the stationary behavior and presents the theoretical framework. For a special case, analytical results are derived that are analogous to the classical solutions for the simple $M/M/C$ queue. The elaborate analysis of a specific case is given to illustrate the basic idea of the framework. A technical proof with respect to the existence of $d-1$ roots is displayed to sustain the corresponding theory.

1. Introduction

The vast majority of queueing models assume a stationary process in order to derive performance measures. In reality, stochastic models whose parameters vary randomly over time depending on the state of some external environment arise naturally in practice. These models play an important role in operations research, applied probability, management science and queueing theory.

The motivation for treating queueing systems in a random environment comes from several fields. In the transportation field, as argued in [2] and [3], the queueing systems equipped with a randomly-varying parameters depending on the state of an external environment can be used to evaluate the influence of incidents on the congestions of vehicles on a part of a highway. Consider a segment of a road subject to traffic incidents. The space occupied by an individual car denotes one queuing server, which starts its service when a car joins the link, the service is completed when the car reaches the end of the link. A roadway section contains hundreds of such servers, so an $M/M/C$ queueing system is a proper approximation. If an incident happens, all the cars on the road reduce their speed until the incident is cleared. In the financial field, the change of bank rate set by the Central bank affects the conditions

* Corresponding author.
E-mail addresses: math_lzm@csu.edu.cn (Z. Liu), yusenlin@csu.edu.cn (S. Yu).
[1] Supported by the National Natural Science Foundation of China (11271373).

http://dx.doi.org/10.1016/j.jmaa.2015.11.074

under which commercial banks give loans to their clients. These, in turn, significantly influence the intensity of clients' arrival. In the biological field, we can think of mRNA strings being transcribed and degraded in a cell, where these transcriptions usually tend to occur in a clustered environment, as argued in [4].

The first systematic study of this field can be found in Yechiali [12]. Yechiali [12] treated an $M/M/1$ system in a two-state Markovian environment. These investigations were continued into the 2000s and 2010s. Boxma and Kurkova [5] discussed an $M/G/1$ queue where the server's speed alternates between two values. Huang and Lee [8] studied the $M/G/1$ queue in a two-state Markovian environment and analyzed the generalized Pollaczek–Khinchin formula. Baykal-Gursoy and Xiao [2] treated an infinite-server queue with service rate varying depending on the state of an alternating renewal process. In [3], the $M/M/C$ systems in a two-state Markovian environment was introduced to model the traffic flow on a roadway link subject to incidents. Mahabhashyam and Gautam [11] analyzed an $M/G/1$ queue where the server's speed varies based on a continuous-time Markov chain. Adan and Kulkarni [1] focused on the systems where the service rate changes randomly for the single server queue. Falin [7] studied $M/M/\infty$ queue in a semi-Markovian environment and obtained the mean queue length. Blom et al. [4] discussed the Markov-modulated infinite-server queue using the stochastic process limit techniques. Cordeiro and Kharoufeh [6] and Kim [10] treated queues in a random environment with some classical policies. Jiang et al. [9] analyzed an $M/G/1$ queue in a random environment subject to disasters interruptions.

The $M/M/1$ queue and $M/M/\infty$ queue in a Markovian environment have been analyzed extensively. But the existing literatures involved in the $M/M/C$ queue in an external environment are relatively rare. The multi-server queueing systems with varying arrival processes and service intensities possess several practical applications. In a multi-product manufacturing network, the arrival rates to different production centers may vary depending on the type of jobs it is processing, environmental conditions and operator experience. Performance characteristics of a multi-server queueing system with varying arrival rates and service rates allow us to analyze the sensitivity of the corresponding model to errors. In a word, the multi-server queue equipped with randomly-varying parameters is a practically important model that deserves a renewed interest.

The rest of this paper is organized as follows. In section 2, we describe the model in detail and present the stability condition. The elaborate proof of a critical theorem is given in section 3. Section 4 is devoted to treating the case when all the traffic intensities are equal. The simplest case when $C = 2$ and $d = 2$ is studied in section 5 and the final expressions are obtained. In section 6, numerical illustrations are added to explain the effect of some parameters on the queue length. Finally, potential applications and further research are presented in section 7.

2. Model description and stability condition

Our paper will focus on an $M/M/C$ queueing system, where the arrivals and service rates are modulated by a random environment for which the underlying process $C(t)$ is an irreducible continuous-time Markov chain on a finite state space $\{1, 2, \cdots, d\}$, d is a positive integer. That is, when the external environment is in state i, the system functions as an $M(\lambda_i)/M(\mu_i)/C$ queue, with Poisson arrival intensity λ_i and service rate μ_i by each server.

To completely describe how the external environment $C(t)$ evolves through time, it will suffice to give a description of both how long it spends in each state and how it makes transitions from one state to the next. The duration of time the external environment $C(t)$ stays in state i is exponentially distributed with parameter q_i. At the end of this holding time, it makes a transition to another state j with probability q_{ij}/q_i, where $j \neq i$. In other words, the infinitesimal generator of the external environment is

$$Q = \begin{pmatrix} -q_1 & q_{12} & \cdots & q_{1d} \\ q_{21} & -q_2 & \cdots & q_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ q_{d1} & q_{d2} & \cdots & -q_d \end{pmatrix}.$$

The stationary distribution of the external environment $C(t)$ is given by $(\pi_1, \pi_2, \cdots, \pi_d)$, being a d-dimensional vector with non-negative entries summing to 1 through solving $(\pi_1, \pi_2, \cdots, \pi_d)Q = \mathbf{0}$. In the following analysis, we assume that $q_{ii} = 0$, $i = 1, 2, \cdots, d$.

A stochastic process $\{C(t), X(t)\}$ describes the system's state at time t as follows: $C(t)$ denotes the state in which the system operates at time t, while $X(t)$ counts the number of customers present in the system at that time. The system is said to be in state (i, m) if $C(t)$ is in state i and there are m customers in the system. The stability condition will be given later, let p_{im} be the limiting probability of the system in state (i, m). That is, $p_{im} = \lim_{t\to\infty} P(C(t) = i, X(t) = m),\ 1 \le i \le d,\ m = 0, 1, 2, \cdots$.

The steady-state balance equations are given below.

For $i = 1, 2, \cdots, d$ and $m = 0$,

$$(\lambda_i + q_i)p_{i0} = \mu_i p_{i1} + \sum_{j=1}^{d} q_{ji} p_{j0}. \tag{1}$$

For $i = 1, 2, \cdots, d$ and $1 \le m \le C - 1$,

$$(\lambda_i + m\mu_i + q_i)p_{im} = \lambda_i p_{i,m-1} + (m+1)\mu_i p_{i,m+1} + \sum_{j=1}^{d} q_{ji} p_{jm}. \tag{2}$$

For $i = 1, 2, \cdots, d$ and $m \ge C$,

$$(\lambda_i + C\mu_i + q_i)p_{im} = \lambda_i p_{i,m-1} + C\mu_i p_{i,m+1} + \sum_{j=1}^{d} q_{ji} p_{jm}. \tag{3}$$

We define $p_{i\cdot} = \sum_{m=0}^{m=\infty} p_{im}$, then $p_{i\cdot}$ is the probability of the external environment being at state i. That is,

$$\pi_i = p_{i\cdot} = \sum_{m=0}^{\infty} p_{im}.$$

Theorem 1. *For the M/M/C queue in an external environment, the stability holds if and only if*

$$\bar{\mu} > \bar{\lambda},$$

where $\bar{\mu} = \sum_{i=1}^{d} C\mu_i \pi_i$, $\bar{\lambda} = \sum_{i=1}^{d} \lambda_i \pi_i$.

Proof. Starting with $m = 0$ and summing each of balance equations (1), (2) and (3) over i, then

$$\sum_{i=1}^{d} \lambda_i p_{im} = \sum_{i=1}^{d} (m+1)\mu_i p_{i,m+1} \quad (0 \le m \le C - 2),$$
$$\sum_{i=1}^{d} \lambda_i p_{im} = \sum_{i=1}^{d} C\mu_i p_{i,m+1} \quad (m \ge C - 1).$$

Summing over all m, we obtain

$$\sum_{i=1}^{d}\lambda_i\pi_i=\sum_{i=1}^{d}C\mu_i(\pi_i-p_{i0})-\sum_{i=1}^{d}\sum_{k=1}^{C-1}(C-k)\mu_i p_{ik}.$$

That is,

$$\sum_{i=1}^{d}\sum_{k=0}^{C-1}(C-k)\mu_i p_{ik}=\sum_{i=1}^{d}C\mu_i\pi_i-\sum_{i=1}^{d}\lambda_i\pi_i=\bar{\mu}-\bar{\lambda}, \tag{4}$$

where $\bar{\mu}=\sum_{i=1}^{d}C\mu_i\pi_i$, $\bar{\lambda}=\sum_{i=1}^{d}\lambda_i\pi_i$.

Note that all the probabilities p_{ik} in (4) can be expressed by p_{i0} $(i=1,2,\cdots,d)$ through solving the balance equations.

In addition, from the theory of recurrent events it can be deduced that the probabilities p_{ik} $(1\le i\le d, k\ge 0)$ are either all positive or, alternatively, all equal to zero. For the stochastic process $\{C(t),X(t)\}$, the steady-state regime exists if and only if $p_{i0}>0$ $(i=1,2,\cdots,d)$. From (4), the necessary and sufficient condition for its existence is $\bar{\mu}>\bar{\lambda}$.

The proof is finished. □

The intuitive interpretation of the theorem is straightforward: note that $\bar{\mu}$ is the average capacity of the system to render service, $\bar{\lambda}$ is the average arrival intensity. For steady-state conditions, the average service capacity must exceed the average arrival rate.

Denote the partial generating functions by

$$G_i(z)=\sum_{m=0}^{\infty}p_{im}z^m,\ |z|\le 1,\ i=1,2,\cdots,d.$$

So the generating function of the steady-state queue length in the system is given by

$$G(z)=\sum_{i=1}^{d}G_i(z).$$

Multiplying both sides of the balance equations by z^m appropriately and summing over all m for state i, then

$$[\lambda_i z(1-z)+q_i z+C\mu_i(z-1)]G_i(z)-\sum_{j=1}^{d}q_{ji}zG_j(z)=(z-1)\sum_{k=0}^{C-1}(C-k)\mu_i z^k p_{ik}. \tag{5}$$

We define

$$f_i(z)=\lambda_i z(1-z)+q_i z+C\mu_i(z-1)\quad (i=1,2,\cdots,d),$$
$$\mathbf{A}(z)=\begin{pmatrix} f_1(z) & -q_{21}z & -q_{31}z & \cdots & -q_{d1}z\\ -q_{12}z & f_2(z) & -q_{32}z & \cdots & -q_{d2}z\\ -q_{13}z & -q_{23}z & f_3(z) & \cdots & -q_{d3}z\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ -q_{1d}z & -q_{2d}z & -q_{3d}z & \ldots & f_d(z)\end{pmatrix},$$

$$\mathbf{g}(z) = \begin{pmatrix} G_1(z) \\ G_2(z) \\ \cdots \\ G_d(z) \end{pmatrix}, \quad \mathbf{b}(z) = \begin{pmatrix} \sum_{k=0}^{C-1}(C-k)\mu_1 z^k p_{1k} \\ \sum_{k=0}^{C-1}(C-k)\mu_2 z^k p_{2k} \\ \cdots \\ \sum_{k=0}^{C-1}(C-k)\mu_d z^k p_{dk} \end{pmatrix}.$$

The equations in (5) can be rewritten as the following matrix equation

$$\mathbf{A}(z)\mathbf{g}(z) = (z-1)\mathbf{b}(z).$$

Applying Cramer's rule, for all values of z at which $\mathbf{A}(z)$ is nonsingular, we have

$$|\mathbf{A}(z)|G_i(z) = |\mathbf{A}_i(z)|(z-1) \ (i = 1, 2, \cdots, d), \tag{6}$$

where $|\mathbf{A}|$ is the determinant of a matrix $\mathbf{A}$ and $\mathbf{A}_i(z)$ is derived through replacing the ith column of $\mathbf{A}(z)$ with $\mathbf{b}(z)$.

All the unknown probabilities in equation (6) can be expressed by p_{i0} $(i = 1, 2, \cdots, d)$ through solving the balance equations, we need to obtain the d unknown probabilities p_{i0} $(i = 1, 2, \cdots, d)$. Note that equation (4) provides a linear relation for p_{i0}. Since $|\mathbf{A}(z)|$ is a polynomial of degree $2d$, we define a new polynomial $Q(z)$ of degree $2d-1$ by

$$|\mathbf{A}(z)| = (z-1)Q(z).$$

Thus

$$G_i(z) = \frac{|\mathbf{A}_i(z)|}{Q(z)}. \tag{7}$$

Clearly, $|\mathbf{A}_i(z)|$ must equal zero whenever $Q(z) = 0$. The following theorem provides a method to obtain the additional $d-1$ linear equations for the d unknown probabilities p_{i0}.

Theorem 2. *The polynomial $Q(z)$ possesses exactly $d-1$ distinct real roots in the interval $(0,1)$.*

The proof of the theorem is fairly delicate and lengthy, we relegate the elaborate proof to the next section.

If we denote the $d-1$ distinct real roots of $Q(z)$ in the interval $(0,1)$ by $z_1, z_2, \cdots, z_{d-1}$, respectively. It immediately follows that

$$|\mathbf{A}_i(z_j)| = 0 \ \ (j = 1, 2, \cdots, d-1; i = 1, 2, \cdots, d). \tag{8}$$

We have arrived at a critical point, the additional $d-1$ linear equations together with the stability condition equation (4) provide d equations. Besides, all the unknown probabilities in the d equations can be expressed by p_{i0}, $i = 1, 2, \cdots, d$. With known p_{i0}, the generating function $G(z)$ can be eventually obtained and the expected number of customers in the system is given by evaluating $G'(z)|_{z=1}$.

In fact, it is difficult to obtain the explicit expressions for the roots z_j $(j = 1, 2, \cdots, d-1)$ and there seems to be no elegant and compact formulas relating unknown probabilities p_{im} in the d equations to p_{i0} $(i = 1, 2, \cdots, d)$, so the closed form results for the generating functions are difficult to obtain, if at all possible. After some technical rearrangement of balance equations, we arrive at a series of recursive formulas:

$$p_{im} = \frac{\lambda_i}{m\mu_i} p_{i,m-1} + \frac{q_i}{m\mu_i} \sum_{k=0}^{m-1} p_{ik} - \sum_{j=1}^{d} \frac{q_{ji}}{m\mu_i} \sum_{k=0}^{m-1} p_{jk} \ \ (1 \le m \le C-1; i = 1, 2, \cdots, d).$$

3. Proof of the theorem

This section is devoted to proving Theorem 2, we introduce a series of polynomials as follows:

$$Q_0(z)=1,\quad Q_1(z)=f_d(z),\quad Q_2(z)=\begin{vmatrix} f_{d-1}(z) & -q_{d,d-1}z \\ -q_{d-1,d}z & f_d(z)\end{vmatrix},$$

$$Q_k(z)=\begin{vmatrix} f_{d-k+1}(z) & -q_{d-k+2,d-k+1}z & \cdots & -q_{d,d-k+1}z \\ -q_{d-k+1,d-k+2}z & f_{d-k+2}(z) & \cdots & -q_{d,d-k+2}z \\ \vdots & \vdots & \ddots & \vdots \\ -q_{d-k+1,d}z & -q_{d-k+2,d}z & \cdots & f_d(z)\end{vmatrix}\ (1\le k\le d-1),$$

$$Q(z)=\frac{|\mathbf{A}(z)|}{z-1}.$$

That is, $Q_k(z)$ $(k=1,2,\cdots,d-1)$ are the determinants of the main-diagonal minors of $\mathbf{A}(z)$ starting from the lower right-hand corner of the matrix. We present several recursive equations, which is crucial for our subsequent analysis.

$$Q_{k+1}(z)=a_k(z)Q_k(z)-b_k(z)Q_{k-1}(z)\ (k=1,2,\cdots,d-2), \tag{9}$$

$$(z-1)Q(z)=c(z)Q_{d-1}(z)-d(z)Q_{d-2}(z). \tag{10}$$

These recursive relations are not difficult to obtain. According to Schur's theorem, if $\mathbf{A}$ and $\mathbf{D}$ are square matrices, $\mathbf{D}$ is invertible, then

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D}\end{vmatrix}=|\mathbf{D}|\cdot|\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}|.$$

In the case that $\mathbf{D}$ is singular, the inverse of $\mathbf{D}$ in the equation can be replaced by a generalized inverse.

We define

$$\mathbf{A}_k(z)=\begin{pmatrix} f_{d-k+1}(z) & -q_{d-k+2,d-k+1}z & \cdots & -q_{d,d-k+1}z \\ -q_{d-k+1,d-k+2}z & f_{d-k+2}(z) & \cdots & -q_{d,d-k+2}z \\ \vdots & \vdots & \ddots & \vdots \\ -q_{d-k+1,d}z & -q_{d-k+2,d}z & \cdots & f_d(z)\end{pmatrix}\ (1\le k\le d-1).$$

That is, $Q_k(z)=|\mathbf{A}_k(z)|=det(\mathbf{A}_k(z))$, $\mathbf{A}_k(z)$ $(k=1,2,\cdots,d-1)$ are the main-diagonal minors of $\mathbf{A}(z)$ starting from the lower right-hand corner of the matrix.

Hence, when $1\le k\le d-2$, we have

$$\begin{aligned} Q_{k+1}(z)&=(|f_{d-k}(z)-\mathbf{B}\mathbf{A}_k^{-1}(z)\mathbf{C}|+1)Q_k(z)-Q_k(z)\\ &=(|f_{d-k}(z)-\mathbf{B}\mathbf{A}_k^{-1}(z)\mathbf{C}|+1)Q_k(z)-|f_{d-k+1}(z)-\mathbf{D}\mathbf{A}_{k-1}^{-1}(z)\mathbf{E}|\cdot Q_{k-1}(z),\end{aligned}$$

and

$$\begin{aligned} (z-1)Q(z)&=(|f_1(z)-\mathbf{F}\mathbf{A}_{d-1}^{-1}(z)\mathbf{G}|+1)Q_{d-1}(z)-Q_{d-1}(z)\\ &=(|f_1(z)-\mathbf{F}\mathbf{A}_{d-1}^{-1}(z)\mathbf{G}|+1)Q_{d-1}(z)-|f_2(z)-\mathbf{H}\mathbf{A}_{d-2}^{-1}(z)\mathbf{J}|\cdot Q_{d-2}(z),\end{aligned}$$

where

$$\begin{aligned}
\mathbf{B} &= (-q_{d-k+1,d-k}z, \cdots, -q_{d,d-k}z) \quad , \quad \mathbf{C} = (-q_{d-k,d-k+1}z, \cdots, -q_{d-k,d}z)^T, \\
\mathbf{D} &= (-q_{d-k+2,d-k+1}z, \cdots, -q_{d,d-k+1}z) \quad , \quad \mathbf{E} = (-q_{d-k+1,d-k+2}z, \cdots, -q_{d-k+1,d}z)^T, \\
\mathbf{F} &= (-q_{21}z, \cdots, -q_{d-1,1}z, -q_{d1}z) \quad , \quad \mathbf{G} = (-q_{12}z, \cdots, -q_{1,d-1}z, -q_{1d}z)^T, \\
\mathbf{H} &= (-q_{32}z, \cdots, -q_{d-1,2}z, -q_{d2}z) \quad , \quad \mathbf{J} = (-q_{23}z, \cdots, -q_{2,d-1}z, -q_{2d}z)^T.
\end{aligned}$$

For these $d+1$ polynomials, we present six properties as follows:

(A) $Q_0(z) = 1$ has no roots in $(0, +\infty)$.

(B) For $k = 1, 2, \cdots, d-2$, $Q_k(z)$ and $Q_{k+1}(z)$ do not possess any joint roots in $(0, +\infty)$. If not, let $z_0 > 0$ be a joint root of $Q_k(z)$ and $Q_{k+1}(z)$, then we have $Q_{k-1}(z_0) = 0$ from equation (9). Similarly, if $Q_k(z_0) = Q_{k-1}(z_0) = 0$, then $Q_{k-2}(z_0) = 0$ from (9). Iterating, we eventually arrive at $Q_0(z_0) = 0$, which contradicts property (A). In an analogous fashion, $Q(z)$ and $Q_{d-1}(z)$ do not possess any joint roots in $(0, +\infty)$.

(C) If $z_0 > 0$, $Q_k(z_0) = 0$ $(k = 1, 2, \cdots, d-2)$. Based on (9), $Q_{k-1}(z_0)$ and $Q_{k+1}(z_0)$ are opposite in sign. If $Q_{d-1}(z_0) = 0$ and $0 < z_0 < 1$, then $Q_{d-2}(z_0)$ and $Q(z_0)$ are the same in sign. If $Q_{d-1}(z_0) = 0$ and $z_0 > 1$, then $Q_{d-2}(z_0)$ is opposite in sign to $Q(z_0)$. These assertions can be verified from (9) and (10).

(D) $Q_k(1) > 0$ $(k = 1, 2, \cdots, d-1)$, $Q(1) > 0$. Setting $z = 1$ in the expressions of $Q_k(z)$, these results can be derived after some manipulations.

(E) $Sign[Q_k(0)] = (-1)^k$, $k = 1, 2, \cdots, d-1$, where $Sign[x] = 1$ if $x > 0$ and $Sign[x] = -1$ if $x < 0$. Note that $f_k(0) = -C\mu_k < 0$, when $1 \le k \le d-1$, $Q_k(0) = (-1)^k \cdot C\mu_{d-k+1} \cdot C\mu_{d-k+2} \cdots C\mu_d$. Besides, $Q(z) = \frac{|\mathbf{A}(z)|}{z-1}$ and $Sign[|\mathbf{A}(0)|] = (-1)^d$, it follows that $Sign[Q(0)] = (-1)^{d+1}$.

(F) $Sign[Q_k(\infty)] = (-1)^k$, $k = 1, 2, \cdots, d-1$, note that the highest-power term of the polynomial $Q_k(z)$ is $(-1)^k \cdot \lambda_{d-k+1} \cdot \lambda_{d-k+2} \cdots \lambda_d z^{2k}$.

To give properties (B) and (C), equations (9) and (10) are needed. We have to point out that when $Q_k(z_0) = Q_{k+1}(z_0) = 0$, equation (9) can be used. The inverses of $\mathbf{A}_k(z_0)$ and $\mathbf{A}_{k-1}(z_0)$ do not exist, but when $\mathbf{A}_k(z_0)$ is singular, the inverse of matrix $\mathbf{A}_k(z_0)$ can be replaced by a generalized inverse, such as the Moore–Penrose pseudoinverse.

The theorem can be proved as follows:

According to (D), (E) and (F), the quadratic polynomial $Q_1(z)$ possesses a root $z_{1,1}$ in $(0, 1)$ and the other root $z_{1,2}$ in $(1, \infty)$. Based on property (C), $Q_2(z_{1,1}) < 0$ and $Q_2(z_{1,2}) < 0$. From (D), (E) and (F) together with the fact that $Q_2(z)$ is a quartic polynomial, we can conclude that each of the intervals $(0, z_{1,1})$, $(z_{1,1}, 1)$, $(1, z_{1,2})$ and $(z_{1,2}, \infty)$ contains exactly one root of $Q_2(z)$. Similarly, each of the intervals $(0, z_{2,1})$, $(z_{2,1}, z_{2,2})$, $(z_{2,2}, 1)$, $(1, z_{2,3})$, $(z_{2,3}, z_{2,4})$ and $(z_{2,4}, \infty)$ contains exactly one root of $Q_3(z)$, and so on.

Repeating this procedure, then $Q_{d-1}(z)$ possesses $2d-2$ real roots in which $d-1$ real roots lie in $(0, 1)$ and $d-1$ roots lie in $(1, \infty)$. The $2d-2$ real roots are denoted, in an increasing order, by $z_{d-1,i}$ $(i = 1, 2, \cdots, 2d-2)$, respectively. According to these properties, we have

$$\begin{aligned}
Sign[Q_{d-2}(z_{d-1,i})] &= (-1)^{d+i-1} \ (i = 1, 2, \cdots, d-1), \\
Sign[Q_{d-2}(z_{d-1,i})] &= (-1)^{d+i} \ (i = d, d+1, \cdots, 2d-2).
\end{aligned}$$

Based on these relations and property (C), we get

$$Sign[Q(z_{d-1,i})] = (-1)^{d+i-1} \ (i = 1, 2, \cdots, 2d-2).$$

Clearly, there is at least one real root of $Q(z)$ between any consecutive roots of $Q_{d-1}(z)$. Since $Sign[Q(0)] = (-1)^{d+1}$ and $Sign[Q(z_{d-1,1})] = (-1)^d$, $Q(z)$ has a root in $(0, z_{d-1,1})$. In addition, $Q(1) > 0$

and $Sign[Q(z_{d-1,d-1})] = (-1)^{2d-2} > 0$, so $Q(z)$ has no root in $(z_{d-1,d-1}, 1)$. Similarly, $Sign[Q(z_{d-1,d})] = (-1)^{2d-1} < 0$, $Q(z)$ possesses a root in $(1, z_{d-1,d})$. Note that $Q(z)$ is a polynomial of degree $2d-1$. It follows that $Q(z)$ possesses exactly $d-1$ real roots in the interval $(0, 1)$.

The proof is finished.

4. A special case: $\lambda_i/\mu_i = \theta$ for all i

Similar to the $M/M/1$ queue and $M/M/\infty$ queue in a multi-phase Markovian random environment, there is one case where a specialized assumption causes the final expressions to be of extreme simplicity. This is the case when all the traffic intensities λ_i/μ_i $(i = 1, 2, \cdots, d)$ are equal, then the system possesses properties similar to a standard $M/M/C$ queue and an explicit simple solution can be derived.

Theorem 3. *If $\lambda_i/\mu_i = \theta$ for all i, $i = 1, 2, \cdots, d$, then*

$$p_{im} = \pi_i p_0 \frac{\theta^m}{m!} \quad (0 \le m \le C-1),$$

$$p_{im} = \pi_i p_0 \frac{(\frac{\theta}{C})^{m-C}\theta^C}{C!} \quad (m \ge C),$$

$$p_0 = [\sum_{k=0}^{C-1} \frac{\theta^k}{k!} + \frac{\theta^C}{C!(1-\frac{\theta}{C})}]^{-1},$$

where $i = 1, 2, \cdots, d$ and $m = 0, 1, 2, \cdots$.

Proof. Clearly, since $(\pi_1, \pi_2, \cdots, \pi_d)Q = \mathbf{0}$, then $q_i\pi_i = \sum_{j=1}^d q_{ji}\pi_j$. When $0 \le m \le C-1$, adding same terms to both sides of the equation $q_i\pi_i = \sum_{j=1}^d q_{ji}\pi_j$, then

$$(\lambda_i + m\mu_i + q_i)\pi_i = m\mu_i\pi_i + \lambda_i\pi_i + \sum_{j=1}^d q_{ji}\pi_j.$$

Multiplying by θ^m and using the assumption that $\lambda_i/\mu_i = \theta$ for all i $(i = 1, 2, \cdots, d)$, we arrive at

$$(\lambda_i + m\mu_i + q_i)\pi_i\theta^m = m\lambda_i\pi_i\theta^{m-1} + \mu_i\pi_i\theta^{m+1} + \sum_{j=1}^d q_{ji}\pi_j\theta^m.$$

Dividing by $m!$ and multiplying by p_0, then

$$(\lambda_i + m\mu_i + q_i)\pi_i p_0 \frac{\theta^m}{m!} = \lambda_i\pi_i p_0 \frac{\theta^{m-1}}{(m-1)!} + (m+1)\mu_i\pi_i p_0 \frac{\theta^{m+1}}{(m+1)!} + \sum_{j=1}^d q_{ji}\pi_j p_0 \frac{\theta^m}{m!}. \tag{11}$$

Setting $p_{im} = \pi_i p_0 \frac{\theta^m}{m!}$ $(0 \le m \le C-1)$ in (11) results in the steady-state balance equations (1) and (2) in section 2. Since these equations possess a unique solution, $p_{im} = \pi_i p_0 \frac{\theta^m}{m!}$ $(0 \le m \le C-1)$ is that solution.

Now, we need to prove the theorem is valid for $m = C$.

According to $(\lambda_i + (C-1)\mu_i + q_i)p_{i,C-1} = \lambda_i p_{i,C-2} + C\mu_i p_{i,C} + \sum_{j=1}^d q_{ji}p_{j,C-1}$ together with the fact that $q_i\pi_i = \sum_{j=1}^d q_{ji}\pi_j$, we get

$$p_{i,C} = p_{i,C-1}(\frac{\theta}{C} + \frac{q_i}{C\mu_i} + \frac{C-1}{C}) - \sum_{j=1}^d \frac{q_{ji}}{C\mu_i} p_{j,C-1} - \frac{\theta}{C} p_{i,C-2}$$

$$
\begin{aligned}
&= p_0 \cdot \frac{\theta^{C-1}}{(C-1)!}[\pi_i(\frac{\theta}{C} + \frac{q_i}{C\mu_i} + \frac{C-1}{C}) - \sum_{j=1}^{d} \frac{q_{ji}}{C\mu_i}\pi_j - \pi_i \frac{C-1}{C}] \\
&= \pi_i p_0 \frac{\theta^C}{C!}.
\end{aligned}
$$

The remaining proof will be by induction. Assume that the theorem holds up to some $m \geq C$, we need to prove it holds for $m+1$ as well. In fact, the procedure is quite similar to the previous proof.

From $(\lambda_i + C\mu_i + q_i)p_{im} = \lambda_i p_{i,m-1} + C\mu_i p_{i,m+1} + \sum_{j=1}^{d} q_{ji} p_{jm}$ $(m \geq C)$, we have

$$
\begin{aligned}
p_{i,m+1} &= p_{im}(\frac{\theta}{C} + \frac{q_i}{C\mu_i} + 1) - \sum_{j=1}^{d} \frac{q_{ji}}{C\mu_i} p_{jm} - \frac{\theta}{C} p_{i,m-1} \\
&= p_0 \cdot \frac{(\frac{\theta}{C})^{m-C}\theta^C}{C!}[\pi_i(\frac{\theta}{C} + \frac{q_i}{C\mu_i} + 1) - \sum_{j=1}^{d} \frac{q_{ji}}{C\mu_i}\pi_j - \pi_i] \\
&= \pi_i p_0 \frac{(\frac{\theta}{C})^{m+1-C}\theta^C}{C!},
\end{aligned}
$$

the expression of p_0 follows from the normalization condition.

The proof is finished. □

5. The case when $C = 2$ and $d = 2$

In this section, we present an example to illustrate the related procedure discussed in section 2. If the number of servers and states are not too large, we can get the explicit expressions after some algebraic operations. When the number of servers and states increase, the related expressions are rather complicated, numerical methods can be employed to solve any specific case. The proof in section 3 suggests also a numerical method for determining the roots.

When $C = 2$ and $d = 2$, $Q(z)$ has a unique root z_0 in $(0, 1)$, equations (4) and (8) can be rewritten as

$$
\begin{gathered}
q_2 z_0(2\mu_2 p_{20} + \mu_2 z_0 p_{21}) + (2\mu_1 p_{10} + \mu_1 z_0 p_{11})[\lambda_2 z_0(1 - z_0) + q_2 z_0 + 2\mu_2(z_0 - 1)] = 0, \\
2\mu_1 p_{10} + \mu_1 p_{11} + 2\mu_2 p_{20} + \mu_2 p_{21} = \bar{\mu} - \bar{\lambda}.
\end{gathered}
$$

The unknown probabilities p_{11} and p_{21} can be eliminated by using balance equations, then

$$
p_{11} = \frac{\lambda_1 + q_1}{\mu_1} p_{10} - \frac{q_2}{\mu_1} p_{20}, \quad p_{21} = \frac{\lambda_2 + q_2}{\mu_2} p_{20} - \frac{q_1}{\mu_2} p_{10}.
$$

After some tedious operations, we arrive at

$$
\begin{gathered}
q_2 z_0[(2\mu_2 + \lambda_2 z_0 + q_2 z_0)p_{20} - q_1 z_0 p_{10}] \\
+ [(2\mu_1 + \lambda_1 z_0 + q_1 z_0)p_{10} - q_2 z_0 p_{20}][\lambda_2 z_0(1 - z_0) + q_2 z_0 + 2\mu_2(z_0 - 1)] = 0, \\
(2\mu_1 + \lambda_1)p_{10} + (2\mu_2 + \lambda_2)p_{20} = \bar{\mu} - \bar{\lambda} = \sum_{i=1}^{2} 2\mu_i \pi_i - \sum_{i=1}^{2} \lambda_i \pi_i,
\end{gathered}
$$

thus

$$
p_{10} = \frac{q_2 z_0(\bar{\mu} - \bar{\lambda})(-\lambda_2 z_0^2 + 2\mu_2 z_0 - 4\mu_2)}{[(2\mu_1 + \lambda_1 z_0 + q_1 z_0)(2\mu_2 + \lambda_2) + q_2 z_0(2\mu_1 + \lambda_1)]f_2(z_0) - g_1(z_0) - q_1 q_2 z_0^2(2\mu_2 + \lambda_2)},
$$

where $f_2(z_0) = \lambda_2 z_0(1 - z_0) + q_2 z_0 + 2\mu_2(z_0 - 1)$, $g_1(z_0) = q_2 z_0(2\mu_1 + \lambda_1)(2\mu_2 + \lambda_2 z_0 + q_2 z_0)$.

In an analogous fashion,

$$p_{20} = \frac{q_1 z_0(\bar{\mu} - \bar{\lambda})(-\lambda_1 z_0^2 + 2\mu_1 z_0 - 4\mu_1)}{[(2\mu_2 + \lambda_2 z_0 + q_2 z_0)(2\mu_1 + \lambda_1) + q_1 z_0(2\mu_2 + \lambda_2)]f_1(z_0) - g_2(z_0) - q_1 q_2 z_0^2(2\mu_1 + \lambda_1)},$$

and $f_1(z_0) = \lambda_1 z_0(1 - z_0) + q_1 z_0 + 2\mu_1(z_0 - 1)$, $g_2(z_0) = q_1 z_0(2\mu_2 + \lambda_2)(2\mu_1 + \lambda_1 z_0 + q_1 z_0)$. Note that the status of state 1 and state 2 are symmetric, so these results are not surprising.

Therefore,

$$G_1(z) = \frac{q_2 z[(2\mu_2 + \lambda_2 z + q_2 z)p_{20} - q_1 z p_{10}] + [(2\mu_1 + \lambda_1 z + q_1 z)p_{10} - q_2 z p_{20}]f_2(z)}{Q(z)},$$

where $f_2(z) = \lambda_2 z(1 - z) + q_2 z + 2\mu_2(z - 1)$.

$$G_2(z) = \frac{q_1 z[(2\mu_1 + \lambda_1 z + q_1 z)p_{10} - q_2 z p_{20}] + [(2\mu_2 + \lambda_2 z + q_2 z)p_{20} - q_1 z p_{10}]f_1(z)}{Q(z)},$$

and $f_1(z) = \lambda_1 z(1 - z) + q_1 z + 2\mu_1(z - 1)$. In this special case, the expression of $Q(z)$ is

$$\begin{aligned} Q(z) =& \lambda_1\lambda_2 z^3 - (\lambda_1\lambda_2 + \lambda_1 q_2 + 2\mu_2\lambda_1 + \lambda_2 q_1 + 2\mu_1\lambda_2)z^2 \\ &+ (2\mu_2\lambda_1 + 2\mu_2 q_1 + 2\mu_1\lambda_2 + 2\mu_1 q_2 + 4\mu_1\mu_2)z - 4\mu_1\mu_2. \end{aligned}$$

Besides, we have

$$\frac{dG_1'(z)}{dz}|_{z=1} = \frac{M_1(2\mu_1 q_2 + 2\mu_2 q_1 - \lambda_1 q_2 - \lambda_2 q_1) - q_2[2\mu_2 p_{20} + \lambda_2 p_{20} + 2\mu_1 p_{10} + \lambda_1 p_{10}] \cdot Q'(1)}{(2\mu_1 q_2 + 2\mu_2 q_1 - \lambda_1 q_2 - \lambda_2 q_1)^2},$$
$$\frac{dG_2'(z)}{dz}|_{z=1} = \frac{M_2(2\mu_1 q_2 + 2\mu_2 q_1 - \lambda_1 q_2 - \lambda_2 q_1) - q_1[2\mu_1 p_{10} + \lambda_1 p_{10} + 2\mu_2 p_{20} + \lambda_2 p_{20}] \cdot Q'(1)}{(2\mu_1 q_2 + 2\mu_2 q_1 - \lambda_1 q_2 - \lambda_2 q_1)^2},$$

where

$$\begin{aligned} M_1 &= 3\lambda_2 q_2 p_{20} + (2\lambda_1 q_2 + 2\mu_1 q_2 + 4\mu_1\mu_2 - 2\lambda_2\mu_1 + 2\lambda_1\mu_2 - \lambda_1\lambda_2 + 2q_1\mu_2 - \lambda_2 q_1)p_{10}, \\ M_2 &= 3\lambda_1 q_1 p_{10} + (2\lambda_2 q_1 + 2\mu_2 q_1 + 4\mu_1\mu_2 - 2\lambda_1\mu_2 + 2\lambda_2\mu_1 - \lambda_1\lambda_2 + 2q_2\mu_1 - \lambda_1 q_2)p_{20}, \\ Q'(1) &= \lim_{z\to 1} Q'(z) = \lambda_1\lambda_2 - 2\lambda_1 q_2 - 2\mu_2\lambda_1 - 2\lambda_2 q_1 - 2\mu_1\lambda_2 + 2\mu_2 q_1 + 2\mu_1 q_2 + 4\mu_1\mu_2. \end{aligned}$$

Hence, the expected number of customers in the system is

$$E(X) = \frac{dG_1'(z)}{dz}|_{z=1} + \frac{dG_2'(z)}{dz}|_{z=1},$$

the final expression of $E(X)$ is

$$\frac{(M_1 + M_2)(2\mu_1 q_2 + 2\mu_2 q_1 - \lambda_1 q_2 - \lambda_2 q_1) - (q_1 + q_2)[2\mu_2 p_{20} + \lambda_2 p_{20} + 2\mu_1 p_{10} + \lambda_1 p_{10}] \cdot Q'(1)}{(2\mu_1 q_2 + 2\mu_2 q_1 - \lambda_1 q_2 - \lambda_2 q_1)^2}.$$

So already one of the simplest queueing models leads to a difficult expression, when the number of servers and states increase we can only expect more complexity. Besides, contrary to vacation queues and $M/M/\infty$ queue in a Markovian random environment discussed in [2], the system we studied does not exhibit the stochastic decomposition property. The $M/M/\infty$ queue in a Markovian random environment has been studied in [2,4,7].

Table 1
Expected number of customers in the system with $\mu_1 = 2\mu_2$, $\lambda_1 = \lambda_2 = 1$, $q_2 = 0.5$.

		$\mu_1 = 0.7$	$\mu_1 = 1.0$	$\mu_1 = 1.3$	$\mu_1 = 1.6$	$\mu_1 = 1.9$	$\mu_1 = 2.2$
$q_1 = 0.4$	p_{10}	0.02729999	0.13693487	0.21001681	0.26193485	0.30058807	0.33042015
	p_{20}	0.01374642	0.07237547	0.11572248	0.14952405	0.17680442	0.19936045
	E(X)	13.04750731	2.42908805	1.44766644	1.05794575	0.84281979	0.70435483
$q_1 = 0.5$	p_{10}	0.01433563	0.11520633	0.18258561	0.23049639	0.26617561	0.29370901
	p_{20}	0.00917322	0.07719050	0.12725731	0.16612122	0.19736451	0.22311605
	E(X)	22.75236290	2.65553983	1.53618829	1.11155841	0.88143138	0.73474338
$q_1 = 0.6$	p_{10}	0.00487983	0.09818415	0.16066035	0.20514359	0.23829220	0.26387885
	p_{20}	0.00380602	0.07999649	0.13588340	0.17910370	0.21373579	0.24220445
	E(X)	61.39558137	2.87415890	1.61507282	1.15784397	0.91421837	0.76029577

Table 2
Expected number of customers in the system with $\mu_1 = 5\mu_2$, $\lambda_1 = \lambda_2 = 1$, $q_2 = 0.5$.

		$\mu_1 = 1.0$	$\mu_1 = 1.3$	$\mu_1 = 1.6$	$\mu_1 = 1.9$	$\mu_1 = 2.2$	$\mu_1 = 2.5$
$q_1 = 0.4$	p_{10}	0.08269894	0.16271610	0.22069180	0.26452250	0.29874196	0.32338640
	p_{20}	0.02913719	0.05906421	0.08251015	0.10180733	0.11827071	0.13217806
	E(X)	5.62680348	2.72388159	1.89175413	1.48141197	1.23077531	1.05363180
$q_1 = 0.5$	p_{10}	0.05521334	0.13035429	0.18498906	0.22638715	0.25875293	0.28468821
	p_{20}	0.02454285	0.05968720	0.08722315	0.10985322	0.12911392	0.14593537
	E(X)	7.95097884	3.22224069	2.14572773	1.64974022	1.35690637	1.16030043
$q_1 = 0.6$	p_{10}	0.03390628	0.10476759	0.15646844	0.19573235	0.22647505	0.25113271
	p_{20}	0.01825278	0.05808634	0.08931032	0.11494483	0.13672059	0.15569278
	E(X)	12.17106966	3.78325539	2.40044081	1.80924337	1.47235279	1.25085869

6. Numerical examples

We derive the explicit expression to the $M/M/2$ queueing system subject to random interruption in section 5. As argued in [2] and [3], $M/M/C$ queueing model under service interruptions could be used to evaluate the traffic flow on a roadway that is subject to incidents. State 1 denotes a normal traffic flow, while state 2 stands for the system is experiencing an interruption. During interruptions, all servers work at lower efficiency. We present numerical examples to illustrate the impact of some parameters on the mean queue length in the system in equilibrium.

Realizing that the stability condition must be satisfied. Table 1 represents the case with minor interruptions, which reduce the service rate to half of its normal value. Table 2 denotes the case with serious interruptions where interruptions cause the service rate to drop to one fifth of the normal service rate.

We can conclude that the number of customers in the system decreases as service rate increases. In both tables, higher q_1 values (i.e., higher incident frequency) lead to more customers in the system. Through comparing Table 1 and Table 2, it can be seen that serious interruptions result in the value of q_1 to have more significant influence on the mean queue length in the system than the minor interruption cases.

7. Conclusions and further research

The research was conducted on the $M/M/C$ queue operating in a Markovian random environment. Generally speaking, the multi-phase generalization of the $M/M/C$ queue will not yield closed-form solutions. For an interesting case when $\lambda_i/\mu_i = \theta$ for all i, simple and elegant results are obtainable. We illustrate the basic idea through analyzing the case when $C = 2$ and $d = 2$, for large C, the associated calculation will be lengthy and we can only expect more complexity.

There are many potential applications of the queueing model studied in this paper. The applications in the field of transportation, finance and manufacturing network have been presented in the introduction.

In modern libraries, the librarians have to check-in and checkout of books manually during the routine maintenance. Our model can evaluate the impact of the special situations.

The $M/M/C$ queue in a semi-Markovian random environment is a direct extension to our study. One can also consider other policies such as the case of random disastrous system's failures, impatient customers and vacations. Another topic is relate our results to an $M/M/\infty$ queue in a random environment discussed in the literatures.

Acknowledgments

The authors would like to thank the referee for his or her careful reading of the manuscript of this paper and his or her helpful suggestions and invaluable comments. This work was supported by the National Natural Science Foundation of China (11271373).

References

[1] I. Adan, V. Kulkarni, Single-server queue with Markov-dependent inter-arrival and service times, Queueing Syst. 45 (2003) 113–134.
[2] M. Baykal-Gursoy, W. Xiao, Stochastic decomposition in M/M/∞ with Markov modulated service rates, Queueing Syst. 48 (2004) 75–88.
[3] M. Baykal-Gursoy, W. Xiao, K. Ozbay, Modeling traffic flow interruped by incidents, European J. Oper. Res. 195 (2009) 127–138.
[4] J. Blom, O. Kella, M. Mandjes, Markov-modulated infinite-server queues with general service times, Queueing Syst. 76 (2014) 403–424.
[5] O. Boxma, I. Kurkova, The M/G/1 queue with two service speeds, Adv. in Appl. Probab. 33 (2001) 520–540.
[6] J. Cordeiro, J. Kharoufeh, The unreliable M/M/1 retrial queue in a random environment, Stoch. Models 28 (1) (2012) 29–48.
[7] G. Falin, The M/M/∞ queue in random environment, Queueing Syst. 58 (2008) 65–76.
[8] L. Huang, T. Lee, Generalized Pollaczek–Khinchin formula for Markov channels, IEEE Trans. Commun. 61 (2013) 3530–3540.
[9] T. Jiang, L. Liu, J. Li, Analysis of the M/G/1 queue in multi-phase random environment with disasters, J. Math. Anal. Appl. 430 (2015) 857–873.
[10] B. Kim, J. Kim, A single server queue with Markov modulated service rates and impatient customers, Perform. Eval. 83–84 (2015) 1–15.
[11] S. Mahabhashyam, N. Gautam, On queues with Markov modulated service rates, Queueing Syst. 51 (2005) 89–113.
[12] U. Yechiali, P. Naor, Queueing problems with heterogeneous arrivals and service, Oper. Res. 19 (1971) 722–734.