ELSEVIER

Operations Research Letters 26 (2000) 67–80

operations research letters

www.elsevier.com/locate/orms

Classification of Markov processes of M/G/1 type with a tree structure and its applications to queueing models

Qi-Ming He *

Department of Industrial Engineering, DalTech, Dalhousie University, P.O. Box 1000, Halifax, Nova Scotia, Canada, B3J 2X4

Received 1 July 1998; received in revised form 1 July 1999

Abstract

This paper studies the classification problem of Markov processes of M/G/1 type with a tree structure. It is shown that the classification of positive recurrence, null recurrence, and transience of the Markov processes of interest is determined completely by the Perron–Frobenius eigenvalue of a nonnegative matrix. The results are used to find classification criteria for a number of discrete time or continuous time queueing systems with multiple types of customers. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Markov process; Queueing theory; Tree structure; Positive recurrence; Null recurrence; Transience; Lyapunov function; Mean drift method

1. Introduction

For many queueing systems with multiple types of customers, their associated queueing processes can be formulated into Markov processes of matrix M/G/1 type or matrix GI/M/1 type with a tree structure. Examples of this sort can be found in [6,16,17]. In [16,17], the M/G/1 paradigm and the GI/M/1 paradigm of Markov processes with a tree structure were studied, respectively. Analytic results were obtained for stationary distributions of states and fundamental periods. But the classification of these Markov processes is still unsolved, i.e., conditions for these Markov processes to be positive recurrent, null recurrent, or transient are yet to be found. As was pointed out in [16,17], the classification problem of these Markov processes is interesting and difficult.

This paper studies the classification problem of the Markov process of M/G/1 type with a tree structure, which is a special case of the Markov process of matrix M/G/1 type with a tree structure introduced in [16]. A simple criterion is found for a complete classification of the Markov process of M/G/1 type with a tree structure. That is: the classification of the Markov process of interest is determined completely by the Perron–Frobenius eigenvalue of a nonnegative matrix. The result is obtained by using the mean-drift

* Fax: +1-902-420-7858.

E-mail address: qi-ming.he@dal.ca (Q.-M. He)

0167-6377/00/$ - see front matter
PII: S0167-6377(99)00057-7

method or Foster's criterion [1,12] and some results from matrix analytic methods [13–16]. Using the results obtained, the classification problem of a number of discrete time or continuous time queueing systems with batch arrivals and multiple types of customers are solved. Generalizations to the Markov process of matrix M/G/1 type with a tree structure are reported in [5].

The work done by Gajrat et al. [2] is closely related to this paper. Gajrat et al. [2] studied a Markov process of random strings which is a hybrid model of the M/G/1 type and the GI/M/1 type Markov processes with a tree structure. They obtained necessary and sufficient conditions for the Markov process to be positive recurrent, null recurrent, or transient, in terms of positive solutions of a finite system of polynomial equations. For the birth-and-death case, they obtained the same simple solution as we obtain in this paper. Malyshey [10,11] studied random strings as Markov processes with a network structure. General classification criteria were obtained and applications to queueing networks were discussed. Differing from the existing work, this paper exploits the special tree structure associated with the Markov process of M/G/1 type and obtains some explicit results.

Results obtained in this paper are used to find conditions to classify some queueing systems with batch arrivals and multiple types of customers in Section 5. Why is it an interesting problem? Is it possible to solve the classification problem of these queueing systems by considering queueing systems with one type of customer? That is: to classify these queueing systems without distinguishing customers of different types. A note after Example 5.2 shows that unless appropriate parameters are chosen such an approach may fail. It also shows how such an appropriate set of parameters can be chosen.

The rest of the paper is organized as follows. In Section 2, a discrete time Markov chain of M/G/1 type with a tree structure is defined. In Section 3, a simple criterion for the classification of the Markov chain of interest is obtained. Section 4 develops simpler criteria for some special cases. In Section 5, criteria for the classification of several queueing systems are developed using the results obtained in Section 3. Finally, in Section 6, the continuous time Markov process of M/G/1 type with a tree structure is studied.

2. Markov chain of M/G/1 type with a tree structure

In this section, a discrete time Markov chain of M/G/1 type with a tree structure is defined. This Markov chain is a special case of the Markov chain of matrix M/G/1 type with a tree structure introduced in [16]. Continuous time Markov processes of M/G/1 type with a tree structure will be defined and discussed briefly in Section 6.

First, we define a K-ary tree. A K-ary tree is a tree in which each node has a parent and K children, except the root node of the tree. The root node of the tree is denoted as 0. Strings of integers between 1 and K are used to represent nodes of the tree. For instance, the kth child of the root node has a representation of k. The jth child of node k has a representation of kj. Let $\aleph = \{J\colon J = k_1k_2\cdots k_n,\ 1 \leqslant k_i \leqslant K,\ 1 \leqslant i \leqslant n,\ n > 0\} \cup \{0\}$. Any string $J \in \aleph$ is a node in the K-ary tree. The length of a string J is defined as the number of integers in the string and is denoted by $|J|$. When $J = 0$, $|J| = 0$. The following two operations related to strings in $\aleph$ are used in this paper.

Addition operation: for $J = k_1 \cdots k_n \in \aleph$ and $H = j_1 \cdots j_t \in \aleph$, $J + H = k_1 \cdots k_n\, j_1 \cdots j_t \in \aleph$.
Subtraction operation: for $J = k_1 \cdots k_n \in \aleph$ and $H = k_t \cdots k_n \in \aleph$, $t > 0$, $J - H = k_1 \cdots k_{t-1} \in \aleph$.

Without loss of generality, no boundary node linked to the root node is considered. The results obtained in this paper can be generalized to cases where boundary nodes do exist.

Consider a discrete time Markov chain $\{X_n,\ n \geqslant 0\}$ for which X_n takes values in $\aleph$. X_n is referred to as the node of the Markov chain at time n. To be called a (homogenous) Markov chain of M/G/1 type with a tree structure, X_n transits at each step to either the parent node of the current node or a descendent of the parent node. All possible transitions and their corresponding probabilities are given as follows. Suppose that $J = k_1 \cdots k_{|J|}$ and $X_n = J + k$. Then the one step transition probabilities are given as:

1. $X_{n+1} = J + k + H$ with probability $a(J + k,\ J + k + H) = a_0(k, H)$ when $H \in \aleph$ and $H > 0$;
2. $X_{n+1} = J + k$ with probability $a(J + k,\ J + k) = a_1(k)$;
3. $X_{n+1} = J + H$ with probability $a(J + k,\ J + H) = a_1(k, H)$ when $H = j_1 \cdots j_{|H|} \in \aleph$ and $0 < j_1 \neq k$;

4. $X_{n+1} = J$ with probability $a(J+k,\ J) = a_2(k)$;
5. $X_{n+1} = H$ with probability $a(0,H) = a_0(0,H)$ when $X_n = 0,\ H \in \aleph$ and $H > 0$; and
6. $X_{n+1} = 0$ with probability $a(0,0) = a_1(0)$ when $X_n = 0$.

Note that transition probabilities depend only on the last integer in the string of the current node. Also note the difference between $a_0(k,H), a_1(k,H), a_1(k)$, and $a_2(k)$. Transition probability $a_1(k,H)$ is introduced since in discrete time queueing systems (Examples 5.1–5.3), arrivals and service completions can occur simultaneously. We set $a_1(k,H) = 0$ when $H = j_1 j_2 \cdots j_n \in \aleph$ and $j_1 = k$ because the transition from $J+k$ to $J+H$ is described by $a_0(k, k+j_2\cdots j_n)$. Transition probabilities satisfy the total law of probability:

$$\sum_{J\in\aleph,\,J>0} [a_0(k,J) + a_1(k,J)] + a_1(k) + a_2(k) = 1, \quad 1 \leqslant k \leqslant K,$$

$$\sum_{J\in\aleph,\,J>0} a_0(0,J) + a_1(0) = 1. \tag{2.1}$$

Let $N(J,k)$ be the number of occurrences of integer k in string J, $1 \leqslant k \leqslant K$. It is easy to see that $|J| = N(J,1) + \cdots + N(J,K)$. Define

$$\lambda(k,j) = \sum_{J\in\aleph,\,J>0} [a_0(k,J) + a_1(k,J)]N(J,j), \quad 0 \leqslant k \leqslant K,\ 1 \leqslant j \leqslant K, \tag{2.2}$$

$$\mu(k) = a_2(k) + \sum_{J\in\aleph,\,J>0} a_1(k,J),\ 1 \leqslant k \leqslant K.$$

Let Λ be a $K \times K$ matrix with the (k,j)th element $\lambda(k,j)$, $1 \leqslant k,\ j \leqslant K$, and M a $K\times K$ matrix with the (k,k)th diagonal element $\mu(k)$ and all other elements zero. We assume that $\mu(k) > 0$, $1 \leqslant k \leqslant K$. Let $P = M^{-1}\Lambda$ and $\mathrm{sp}(P)$ be the Perron–Frobenius eigenvalue of the nonnegative matrix P (i.e., the eigenvalue with the largest modulus). We call $\mathrm{sp}(P)$ the transition intensity.

Matrix Λ represents the transition rates that the Markov chain will move from the current node to a descendent of its parent. Matrix M represents the transition rates that the Markov chain will move from the current node to its parent. Suppose that an integer k in string $J \in \aleph$ represents a type k customer. Then each element of matrix Λ is the average number of customers who arrive in a unit period of time and each element of matrix M is the average number of customers served in a unit period of time. Matrix P is the ratio of arrival rates to service rates. Intuitively, the classification of Markov chain $\{X_n,\ n\geqslant 0\}$ is determined by matrix P in a way similar to the classification of random walks on nonnegative integers [8]. The objective of this paper is to show that this intuition is true and the classification of Markov chain $\{X_n,\ n\geqslant 0\}$ is determined completely by the Perron–Frobenius eigenvalue $\mathrm{sp}(P)$ of matrix P.

Throughout this paper, it is assumed that Markov chain $\{X_n,\ n\geqslant 0\}$ is irreducible, $\{\lambda(k,j),\ 0\leqslant k\leqslant K,\ 1\leqslant j\leqslant K\}$ are finite, and matrix P is irreducible. Note that when matrix P is reducible, the set $\{1,2,\ldots,K\}$ can be decomposed into disconnected subsets each with an irreducible subset. Results obtained in this paper apply to these irreducible subsets. These assumptions are not restrictive since they are satisfied in many applications.

3. The main theorem

In this section, a complete classification of the Markov chain $\{X_n,\ n\geqslant 0\}$ defined in Section 2 is obtained by using the mean-drift method or Foster's criterion and matrix analytic methods.

Let R_+ be the set of nonnegative real numbers and $R_+^K = R_+ \times \cdots \times R_+$. For vector $\boldsymbol{x} = (x_1,\ldots,x_K)$, its norm is defined as $\|\boldsymbol{x}\| = |x_1| + \cdots + |x_K|$. Note that $(\boldsymbol{x})_i$ represents the ith element of vector $\boldsymbol{x}$. Let

$$\eta \equiv \min_{\{\boldsymbol{b}\in R_+^K,\ \|\boldsymbol{b}\|=1\}} \left\{ \max_{1\leqslant k\leqslant K} \{((\Lambda - M)\boldsymbol{b})_k\} \right\}. \tag{3.1}$$

Note. The set of vector $\boldsymbol{b}$ in Eq. (3.1) can be chosen differently. For instance, set $\{\boldsymbol{b}\colon \boldsymbol{b}\in R_+^K$ and $c_1 \leqslant \|\boldsymbol{b}\| \leqslant c_2\}$ with $0 < c_1 \leqslant c_2 < \infty$ can be used for the same purpose. Also the norm of vector $\boldsymbol{x}$ can be defined differently. For instance, $\|\boldsymbol{x}\| = (x_1^2 + \cdots + x_K^2)^{0.5}$ can be used.

We shall prove that the transition intensity $\mathrm{sp}(P)$, or equivalently η, gives a complete classification of the Markov chain defined in Section 2. The following lemma shows a relationship between η and $\mathrm{sp}(P)$.

Lemma 3.1. *For the Markov chain $\{X_n,\ n\geqslant 0\}$ defined in Section 2, a relationship between η and* sp(P) *is*
(a) sp$(P) < 1$ *if and only if* $\eta < 0$;
(b) sp$(P) = 1$ *if and only if* $\eta = 0$;
(c) sp$(P) > 1$ *if and only if* $\eta > 0$.

Proof. Denote by $\boldsymbol{u}$ a right eigenvector of P associated with the Perron–Frobenius eigenvalue sp(P), i.e., $P\boldsymbol{u} = \mathrm{sp}(P)\boldsymbol{u}$ and $\boldsymbol{u} \neq 0$. Since P is nonnegative and irreducible, every element of vector $\boldsymbol{u}$ is positive [3]. By the definition of matrix P, it has $\Lambda\boldsymbol{u} = \mathrm{sp}(P)M\boldsymbol{u}$. This leads to $(\Lambda - M)\boldsymbol{u} = (\mathrm{sp}(P) - 1)M\boldsymbol{u}$. Note that every element of vector $M\boldsymbol{u}$ is positive.

Suppose that sp$(P) < 1$. Then $(\mathrm{sp}(P) - 1)M\boldsymbol{u} < 0$. Then

$$\begin{aligned}\eta &\leqslant \max_{1\leqslant k\leqslant K}\{((\Lambda - M)\boldsymbol{u})_k\}\\ &= \max_{1\leqslant k\leqslant K}\{(\mathrm{sp}(P) - 1)M\boldsymbol{u})_k\} < 0.\end{aligned}$$

Suppose that $\eta < 0$. Then there exists a vector $\boldsymbol{b}^* = (b_1^*,\ldots,b_K^*)^{\mathrm{T}}$ with $b_k^* \geqslant 0$ and $\|\ \boldsymbol{b}^*\ \| = 1$ such that $(\Lambda - M)\boldsymbol{b}^* \leqslant \eta\boldsymbol{e}$, where "T" represents the transpose of matrix and $\boldsymbol{e}$ is the vector with all elements one, which implies $\Lambda\boldsymbol{b}^* < M\boldsymbol{b}^*$ and $P\boldsymbol{b}^* < \boldsymbol{b}^*$. Let $\boldsymbol{y}$ be the positive left eigenvector corresponding to sp(P). Then we have $\boldsymbol{y}P\boldsymbol{b}^* = \mathrm{sp}(P)\boldsymbol{y}\boldsymbol{b}^* < \boldsymbol{y}\boldsymbol{b}^*$. Since $\boldsymbol{y}$ is positive and $\boldsymbol{b}^*$ is nonnegative and nonzero, we must have sp$(P) < 1$. This proves Part (a).

Next, we prove Part (c). Suppose that sp$(P) > 1$. If $\eta \leqslant 0$, there exists a nonzero nonnegative vector $\boldsymbol{b}$ such that $(\Lambda - M)\boldsymbol{b} \leqslant 0$, which implies that $P\boldsymbol{b} \leqslant \boldsymbol{b}$. Let $\boldsymbol{y}$ be the positive left eigenvector corresponding to sp(P). Then $\boldsymbol{y}P\boldsymbol{b} \leqslant \boldsymbol{y}\boldsymbol{b}$ implies that sp$(P) \leqslant 1$, which is a contradiction. Therefore, $\eta > 0$. Suppose that $\eta > 0$, but sp$(P) \leqslant 1$. Then $(\Lambda - M)\boldsymbol{u} \leqslant 0$, which implies that $\eta \leqslant 0$. This is a contradiction. Therefore, sp$(P) > 1$. This proves (c).

Part (b) is obtained from (a) and (c). This completes the proof. □

Now, we are ready to state and prove the main theorem of this paper.

Theorem 3.2. *For the Markov chain of* M/G/1 *type with a tree structure — $\{X_n,\ n\geqslant 0\}$ — defined in Section* 2, *it is*
(1) *positive recurrent if and only if* sp$(P) < 1$ *or equivalently* $\eta < 0$;
(2) *null recurrent if and only if* sp$(P) = 1$ *or equivalently* $\eta = 0$;
(3) *transient if and only if* sp$(P) > 1$ *or equivalently* $\eta > 0$.

Proof. The following proof consists of two parts. First, we prove that $\eta < 0$ is a necessary and sufficient condition for positive recurrence of the Markov chain of interest, i.e., Part (1). Then we prove Part (3). As a result of Parts (1) and (3), Part (2) is obtained.

We begin with Part (1). To prove that $\eta < 0$ is sufficient for positive recurrence of the Markov chain of interest, the mean-drift method is applied. The idea of the mean-drift method is to find a Lyapunov function (or test function) $f(J)$ defined on $\aleph$ such that $f(J) \to \infty$ when $|J| \to \infty$ and

$$\begin{aligned}&\boldsymbol{E}[f(X_{n+1}) - f(X_n)\,|\,X_n = J]\\ &\quad= \sum_{H\in\aleph} a(J,H)f(H) - f(J) < -\varepsilon\end{aligned} \tag{3.2}$$

holds for all but a finite number of states in $\aleph$ for some positive ε. If so, the Markov chain is positive recurrent [12]. Suppose that $\eta < 0$. Then there exists a vector $\boldsymbol{b}^* = (b_1^*,\ldots,b_K^*)^{\mathrm{T}}$ with $b_k^* \geqslant 0$ and $\|\ \boldsymbol{b}^*\ \| = 1$ such that $(\Lambda - M)\boldsymbol{b}^* \leqslant \eta\boldsymbol{e}$. When real number δ is small enough, we have $(\Lambda - M)(\boldsymbol{b}^* + \delta\boldsymbol{e}) \leqslant \eta\boldsymbol{e} + \delta(\Lambda - M)\boldsymbol{e} < \eta\boldsymbol{e}/2$. Choose $\boldsymbol{b} = (b_1,\ldots,b_K) = \boldsymbol{b}^* + \delta\boldsymbol{e}$ with a small and positive δ. Then every element of vector $\boldsymbol{b}$ is positive and every element of vector $(\Lambda - M)\boldsymbol{b}$ is less than $\eta/2$. Choose $\varepsilon = \eta/2$. Then $(\Lambda - M)\boldsymbol{b} < -\varepsilon\boldsymbol{e}$ holds. Based on the tree structure of the Markov chain of interest, the following Lyapunov function is introduced. For $J \in \aleph$, define

$$f(J) = \sum_{k=1}^{K} N(J,k)b_k. \tag{3.3}$$

Since every element of vector $\boldsymbol{b}$ is positive, $f(J) \to \infty$ when $|J| \to \infty$. It is clear that function $f(J)$ is additive, i.e., $f(J + H) = f(J) + f(H)$ for $J, H \in \aleph$. Applying the above Lyapunov function, the left-hand side of inequality (3.2) becomes, for $J = k_1 \cdots k_n$ and $k > 0$,

$$\begin{aligned}\sum_{H\in\aleph,\, H>0} &[a_0(k,H)f(J + k + H)\\ &+ a_1(k,H)f(J + H)]\end{aligned}$$

$$
\begin{aligned}
&+a_1(k)f(J+k)+a_2(k)f(J)-f(J+k)\\
&=\sum_{H\in\aleph,\,H>0}\{a_0(k,H)[f(J+k+H)-f(J+k)]\\
&\quad+a_1(k,H)[f(J+H)-f(J+k)]\}\\
&\quad+a_2(k)[f(J)-f(J+k)]\\
&=\sum_{H\in\aleph,\,H>0}\{a_0(k,H)f(H)+a_1(k,H)[f(H)\\
&\quad-f(k)]\}-a_2(k)f(k)\\
&=\sum_{H\in\aleph,\,H>0}[a_0(k,H)f(H)+a_1(k,H)f(H)]\\
&\quad-\left[a_2(k)f(k)+\sum_{H\in\aleph,\,H>0}a_1(k,H)f(k)\right]\\
&=\sum_{H\in\aleph,\,H>0}[a_0(k,H)+a_1(k,H)]\\
&\quad\times\left(\sum_{j=1}^{K}N(H,j)b_j\right)-\mu(k)b_k\\
&=\sum_{j=1}^{K}\lambda(k,j)b_j-\mu(k)b_k\\
&=((\Lambda-M)\boldsymbol{b})_k<-\varepsilon.
\end{aligned}
\tag{3.4}
$$

In the above evaluation, equalities in Eqs. (2.1) and (2.2) are used. Thus, inequality (3.2) holds for all $J\in\aleph$ but $J=0$. By assumptions made in Section 2, $\{\lambda(0,j), 1\leqslant j\leqslant K\}$ are finite, which implies that the left-hand side of inequality (3.2) is finite when $J=0$. Therefore, the Markov chain is positive recurrent.

To prove the necessity of $\eta<0$ for positive recurrence, denote by $v(J)$ the mean first passage time from node J to root node 0 ($v(0)=0$). According to Foster's criterion, when the Markov chain is positive recurrent, $\{v(J),\ J\in\aleph\}$ are finite and satisfy

$$
v(J)=1+\sum_{H\in\aleph}a(J,H)v(H). \tag{3.5}
$$

The above equality leads to $0=1+\sum_H a(J,H)[v(H)v(J)]$ by using the law of total probability. Because of the tree structure and the transition pattern of the Markov chain, function $v(J)$ is an additive function, i.e., $v(J+H)=v(J)+v(H)$, which implies that $v(J)=\sum_{k=1}^{K}N(J,k)v(k)$. This leads to, for $J=k_1\cdots k_n$ and $k>0$,

$$
\begin{aligned}
-1&=\sum_{H\in\aleph,\,H>0}\{a_0(k,H)[v(J+k+H)-v(J+k)]\\
&\quad+a_1(k,H)[v(J+H)-v(J+k)]\}\\
&\quad+a_2(k)[v(J)-v(J+k)]\\
&=\sum_{H\in\aleph,\,H>0}\{a_0(k,H)v(H)+a_1(k,H)\\
&\quad\times[v(H)-v(k)]\}-a_2(k)v(k)\\
&=\sum_{j=1}^{K}\lambda(k,j)v(j)-\mu(k)v(k).
\end{aligned}
\tag{3.6}
$$

Let $\boldsymbol{v}=(v(1),\ldots,v(K))^{\mathrm{T}}$. Eq. (3.6) leads to $(\Lambda-M)\boldsymbol{v}=\boldsymbol{e}$. Since $v(k)\geqslant 1$, $1\leqslant k\leqslant K$, i.e., $\boldsymbol{v}\neq 0$, we obtain $\eta<0$. This completes the proof of Part (1).

We now prove Part (3), i.e., the Markov chain is transient if and only if $\eta>0$. The idea is to look at the probability that the Markov chain will reach the root node from any other node eventually. When the Markov chain is transient, the probability should be less than one. Define sequences $\{G(J)[n],\ n\geqslant 0\}$ for all $J\in\aleph$ as follows. Let $G(0)[n]=1$, for $n\geqslant 0$. Let $G(J)[0]=0$ for $J\in\aleph$ and $J>0$, and

$$
G(J)[n+1]=\sum_{H\in\aleph}a(J,H)G(H)[n]. \tag{3.7}
$$

It is easy to see that $\{G(J)[n],\ n\geqslant 0\}$ is a uniformly bounded and nondecreasing sequence for each $J\in\aleph$. In addition, $G(J)[n]$ is the probability that the Markov chain reaches the root node within n transitions, given that the Markov chain is in node J initially. Note that the root node has become an absorption node. Denote by $G(J)$ the limit of $\{G(J)[n],\ n\geqslant 0\}$ for $J\in\aleph$. It has been proven in [16] that $G(J)$ is the conditional probability that the Markov chain will eventually reach the root node, given that the Markov chain is initially in node J, for each $J\in\aleph$.

When the Markov chain is transient, since matrix P is irreducible, $G(J)<1$ for all $J\in\aleph$ and $J>0$. Denote by $g(k)=1-G(k)$ (>0), $1\leqslant k\leqslant K$, and $g(0)=0$. Because of the special tree structure, it is easy to see that $G(J)=G(j_{|J|})\cdots G(j_1)$ for $J=j_1\cdots j_{|J|}\in\aleph$, which leads to $G(J)=(1-g(j_{|J|}))\cdots(1-g(j_1))\geqslant 1-[g(j_1)+\cdots+g(j_{|J|})]$.

Use the inequality to obtain

$$1 - g(k) = G(k) = \sum_{J:J\in\aleph} a(k,J)G(J)$$

$$\geqslant \sum_{J:J\in\aleph} a(k,J)\{1 - [(g(j_1) + \cdots + g(j_{|J|})]\}$$

$$\Rightarrow 0 \leqslant \sum_{j=1}^{K} \lambda(k,j)g(j) - \mu(k)g(k)$$

$$= ((\Lambda - M)\boldsymbol{g})_k, \tag{3.8}$$

where $\boldsymbol{g} = (g(1),\ldots,g(K))^{\mathrm{T}}$. If $(\Lambda - M)\boldsymbol{g} = 0$, then it can be shown that the Markov chain is recurrent by using the mean-drift method and the Lyapunov function defined in Eq. (3.3) with the positive vector $\boldsymbol{g}$. However, the Markov chain is transient. Therefore, we must have $(\Lambda - M)\boldsymbol{g} \geqslant 0$ and $(\Lambda - M)\boldsymbol{g} \neq 0$, which implies $\mathrm{sp}(P) > 1$ and $\eta > 0$.

Now suppose that $\eta > 0$. We shall show that at least one sequence $\{G(J)[n],\ n\geqslant 0\}$, for $J \in \aleph$, does not converge to one. First, let $p(k)[n]$ be the probability that the Markov chain reaches node J for the first time at the nth transition, given that the Markov chain is in node $J + k$ initially. Then for $J = j_1 \cdots j_{|J|}$,

$$G(J)[n] = \sum_{t=1}^{n-1} p(j_{|J|})[t]G(J - j_{|J|})[n - t]$$

$$\leqslant \left(\sum_{t=1}^{n-1} p(j_{|J|})[t]\right) G(J - j_{|J|})[n - 1]$$

$$\leqslant G(j_{|J|})[n]G(J - j_{|J|})[n] \leqslant \cdots$$

$$\leqslant \prod_{t=|J|}^{1} G(j_t)[n]. \tag{3.9}$$

In the above equation, the fact that $\{G(J)[n],\ n\geqslant 0\}$ is nondecreasing with respect to n is used. Let $g(k)[n] = 1 - G(k)[n],\ 1 \leqslant k \leqslant K$. Since $\{G(k)[0] = 0,\ 1 \leqslant k \leqslant K\}$, it can be proved by induction that $G(k)[n] < 1$, which implies $\{g(k)[n] > 0,\ n > 0\}$, for at least one k ($1 \leqslant k \leqslant K$). Then for at least one k, Eq. (3.7) becomes

$$G(k)[n + 1] = \sum_{J\in\aleph} a(k,J)G(J)[n]$$

$$\leqslant \sum_{J\in\aleph} a(k,J) \prod_{t=|J|}^{1} G(j_t)[n]$$

$$= \sum_{J\in\aleph} a(k,J) \prod_{t=|J|}^{1} (1 - g(j_t)[n])$$

$$= 1 - \sum_{J\in\aleph} a(k,J) \left(\sum_{t=|J|}^{1} g(j_t)[n]\right) + \mathrm{O}\left(\left(\max_{1\leqslant k\leqslant K} \{g(k)[n]\}\right)^2\right)$$

$$= 1 - g(k)[n] - \left[\sum_{j=1}^{K} \lambda(k,j)g(j)[n] - \mu(k)g(k)[n]\right] + \mathrm{O}\left(\left(\max_{1\leqslant k\leqslant K} \{g(k)[n]\}\right)^2\right)$$

$$< G(k)[n] - 0.5\eta \parallel \boldsymbol{g}[n] \parallel + \mathrm{O}(K \parallel \boldsymbol{g}[n] \parallel^2), \tag{3.10}$$

where $\boldsymbol{g}[n] = (g(1)[n],\ldots,g(K)[n])^{\mathrm{T}} \in R_+^K$ and $\boldsymbol{g}[n] \neq 0$. Since $\eta > 0$, there is at least one element of vector $(\Lambda - M)\boldsymbol{g}[n]$ that is positive. The last inequality in equality (3.10) holds for at least one k.

Suppose that all sequences $\{G(J)[n],\ n\geqslant 0\}$, for $J \in \aleph$, converge to one. Then $\{g(k)[n],\ n\geqslant 0\}$ converges to zero uniformly for $1 \leqslant k \leqslant K$. When n is large enough or equivalently $\{g(k)[n],\ 1 \leqslant k \leqslant K\}$ are small enough, inequality (3.10) implies that, for at least one k, $G(k)[n + 1] < G(k)[n]$, which contradicts the fact that the sequence $\{G(k)[n]\}$ is nondecreasing. Therefore, at least one $G(k)$ is less than one, i.e., the Markov chain is transient. This completes the proof of Part (c).

Part (b) holds since Parts (a) and (c) are true. This completes the proof of the theorem. □

Note. When the one-step transition of the Markov chain of interest is constrained to the parent of the current node or children and grandchildren of the parent node (the birth-and-death case), Theorem 3.2 has been proved in [2]. The method used in Theorem 3.2 to prove that $\mathrm{sp}(P) < 1$ is a necessary and sufficient condition for positive recurrence of Markov chain $\{X_n,\ n\geqslant 0\}$ is similar to that of Theorem 2.1 in [2]. But the method used in Theorem 3.2 to prove that $\mathrm{sp}(P) > 1$ is a necessary and sufficient condition for transience of Markov chain $\{X_n,\ n\geqslant 0\}$ is different from the one used in [2]. In fact, the method used in [2] to prove the necessary and sufficient condition for transience requires that $a(J,H)=0$ if $\|\,H|-|J\|\ >d$ for some integer d. Thus, that method cannot be used directly in the M/G/1 case. Besides, the condition $a(J,H)=0$ if $\|\,H|-|J\,\|\ >d$ for some integer d is restrictive when we study Markov processes associated with queueing systems of M/G/1 type.

Theorem 3.2 shows that the classification of the Markov chain of interest is completely determined by transition intensity sp(P) or the value of η. Therefore, to find whether the Markov chain is positive recurrent, null recurrent, or transient, we only need to find the eigenvalue sp(P) or the solution of Eq. (3.1) η. There are many existing methods to find the Perron–Frobenius eigenvalue of a nonnegative matrix, for instance, Elsner's algorithm (Chapter 1 of Neuts [14]). The computation of η can be done by transforming Eq. (3.1) into the following linear programming problem:

$$\begin{aligned}\eta= &\min\ \{x\}\\ &\text{s.t.}\quad (\Lambda-M)\boldsymbol{b}\leqslant x\boldsymbol{e}, \qquad (3.11)\\ &\qquad\ \boldsymbol{be}=1,\ \boldsymbol{b}\in\mathbb{R}_+^K;\ -\infty<x<\infty.\end{aligned}$$

The solution of the above linear programming problem provides information for a complete classification of the Markov chain defined in Section 2. Similar results have been obtained in the theory of queueing networks. For instance, Kumar and Meyn [9] have proved that the solution of a linear programming problem provides information for the stability of reentry network queueing systems.

For Markov chains on a tree structure, the simple rule used to determine whether or not a simple random walk on nonnegative integers is positive recurrent is no longer useful. The reason is that there are several dimensions and the transition rates on one direction may fail to predict positive recurrence of the Markov chain. Surprisingly, Lemma 3.1 and Theorem 3.2 show that the classification of the Markov chain of interest is determined completely by the projections of arrival rates Λ and service rates M on a special direction in R_+^K. This special direction is the eigenvector $\boldsymbol{b}$ corresponding to the Perron–Frobenius eigenvalue of nonnegative matrix P. The Markov chain is positive recurrent if and only if the projection of the arrival rates Λ of every type of customer on direction $\boldsymbol{b}$ — $(\Lambda\boldsymbol{b})_k$ — is smaller than the projection of service rates M of that type of customer — $(M\boldsymbol{b})_k$. Similar interpretations go to the null recurrent and transient cases. In addition, Eq. (3.1) shows that when the Markov chain is transient, for any direction in R_+^K, there exists at least one type of customer whose projection of arrival rates on that direction is larger than that of their service rates.

For queueing applications, we introduce the following decomposition technique. We decompose probabilities $a_0(k,H)$ and $a_1(k)$ into

$$\begin{aligned}&a_0(k,H)=\hat{a}_0(k,H)+\hat{a}_1(k,k+H),\\ &\qquad H\in\aleph,\ H>0, \qquad (3.12)\\ &a_1(k)=\hat{a}_1(k)+\hat{a}_1(k,k).\end{aligned}$$

Note that there is no restriction on this decomposition except that all the numbers must be nonnegative. An interpretation of such a decomposition is that the event {transition from $J+k$ to $J+k+H$} can be decomposed into two disjoint events with probabilities $\hat{a}_0(k,H)$ and $\hat{a}_1(k,k+H)$, respectively. Let $\hat{a}_2(k)=a_2(k)$ and $\hat{a}_1(k,H)=a_1(k,H)$ when $H=j_1\cdots j_{|H|}\in\aleph$ and $0<j_1\neq k$. Transition probabilities $\{\hat{a}_0(k,H),\hat{a}_1(k,H),\hat{a}_1(k),\hat{a}_2(k),\ H\in\aleph,\ H>0\}$ define the same Markov chain in understanding that the transition probabilities from $J+k$ to $J+k+H$ $(H\in\aleph)$ are given by Eq. (3.12). Define

$$\begin{aligned}&\hat{\lambda}(k,j)=\sum_{J\in\aleph,J>0}[\hat{a}_0(k,J)+\hat{a}_1(k,J)]N(J,j),\\ &\qquad 1\leqslant k\leqslant K,1\leqslant j\leqslant K, \qquad (3.13)\\ &\hat{\mu}(k)=\hat{a}_2(k)+\sum_{J\in\aleph,J>0}\hat{a}_1(k,J),\quad 1\leqslant k\leqslant K.\end{aligned}$$

Let $\hat{\Lambda}$ be a $K\times K$ matrix with the (k,j)th element $\hat{\lambda}(k,j)$, $1\leqslant k,\ j\leqslant K$, $\hat{M}$ a $K\times K$ matrix with the (k,k)th diagonal element $\hat{\mu}(k)$ and all other elements zero, $\hat{P}=\hat{M}^{-1}\hat{\Lambda}$. Then the following relationships hold:

$$\begin{aligned}&\hat{\lambda}(k,j)=\lambda(k,j)\quad\text{for } k\neq j;\\ &\hat{\lambda}(k,k)=\lambda(k,k)+\sum_{J\in\aleph}\hat{a}_1(k,k+J),\\ &\hat{\mu}(k)=\mu(k)+\sum_{J\in\aleph}\hat{a}_1(k,k+J), \qquad (3.14)\end{aligned}$$

$\hat{\Lambda} = \Lambda + \Delta$; $\hat{M} = M + \Delta$, where

$$\Delta = \begin{bmatrix} \sum_{J\in\aleph} \hat{a}_1(1,1+J) & & \\ & \ddots & \\ & & \sum_{J\in\aleph} \hat{a}_1(K,K+J) \end{bmatrix}.$$

Define $\hat{\eta}$ by Eq. (3.1) with matrices $\hat{M}$ and $\hat{\Lambda}$. Since $\hat{M} - \hat{\Lambda} = M - \Lambda$, we obtain $\hat{\eta} = \eta$ (of M and Λ). It is easy to prove that $\mathrm{sp}(P) \leqslant \mathrm{sp}(\hat{P}) < 1$ when $\mathrm{sp}(P) < 1$; $\mathrm{sp}(P) = \mathrm{sp}(\hat{P}) = 1$ when $\mathrm{sp}(P) = 1$; and $\mathrm{sp}(P) \geqslant \mathrm{sp}(\hat{P}) > 1$ when $\mathrm{sp}(P) > 1$. Thus, we have proved the following result.

Theorem 3.3. *For $\hat{\eta}$ and $\mathrm{sp}(\hat{P})$ associated with matrices $\hat{M}$ and $\hat{\Lambda}$, conclusions in Lemma* 3.1 *and Theorem* 3.2 *hold.*

By Theorem 3.3, instead of looking for η or $\mathrm{sp}(P)$, we can try to find $\hat{\eta}$ or $\mathrm{sp}(\hat{P})$ to determine whether the Markov chain of interest is positive recurrent, null recurrent, or transient. This technique proves useful in solving the classification problem of some queueing systems where a way of decomposition arises naturally (Examples 5.1–5.4). For these cases, $\{\hat{\lambda}(k,j)\}(\hat{\mu}(k))$, not $\{\lambda(k,j)\}(\mu(k))$, truly represents the arrival (service) rates of customers; $\mathrm{sp}(\hat{P})$, not $\mathrm{sp}(P)$, is equal to the traffic intensity. To avoid heavy notation, we shall only use notation without "^" when the definition of the Markov chain is clear.

4. Some interesting cases

In this section, we use some examples to show when the classification of Markov chain $\{X_n,\ n\geqslant 0\}$ can be easier. Basically, we want to know when $\mathrm{sp}(P)$ is less than, equal to, or larger than one without calculating the Perron–Frobenius eigenvalue of P or solving linear programming problem (3.11). Let

$$\rho(k) = \sum_{j=1}^{K} \frac{\lambda(k,j)}{\mu(j)}, \quad 1\leqslant k\leqslant K. \tag{4.1}$$

Intuitively, $\{\rho(k),\ 1\leqslant k\leqslant K\}$ have much to do with the positive recurrence of Markov chain $\{X_n,\ n\geqslant 0\}$. In fact, the following conclusion holds.

Corollary 4.1. *The Markov chain defined in Section* 2 *is positive recurrent if $\rho(k) < 1$, $1\leqslant k\leqslant K$.*

Proof. By Theorem 3.2, we only have to prove that there exists a positive vector $\boldsymbol{b}$ such that $(\Lambda - M)\boldsymbol{b} < 0$. Let $b_j = 1/\mu(j)$, $1\leqslant j\leqslant K$. It is easy to show that $\max_{1\leqslant k\leqslant K} \{((\Lambda - M)\boldsymbol{b})_k\} = \max_{1\leqslant k\leqslant K} \{\rho(k) - 1\} < 0$. Therefore, $\eta < 0$. Notice that vector $\boldsymbol{b}$ should be normalized to have $\|\boldsymbol{b}\| = 1$, but it is not essential. Thus, the Markov chain is positive recurrent. This completes the proof. □

Corollary 4.1 shows that $\mathrm{sp}(P) \leqslant \max_{1\leqslant k\leqslant K} \{\rho(k)\}$. But condition $\rho(k) < 1$, $1\leqslant k\leqslant K$, is unnecessary for positive recurrence of the Markov chain. The vector $\boldsymbol{b}$ introduced in the proof of Corollary 4.1 is a special direction, which may or may not be the eigenvector corresponding to the Perron–Frobenius eigenvalue of matrix P. Failure of $(\Lambda - M)\boldsymbol{b} < 0$ in this particular direction does not imply that the Markov chain is not positive recurrent. Here is an example.

Example 4.1. Consider a Markov chain of M/G/1 type with a tree structure and $K = 2$. Transition probabilities (only nonzero ones) of this Markov chain are given as

$a_0(1,1) = 0.1$, $a_0(1,2) = 0.2$,

$a_0(2,1) = 0.1$, $a_0(2,2) = 0.1$,

$a_1(2) = 0.3$, $a_1(1,2) = 0.3$,

$a_2(1) = 0.4$, $a_2(2) = 0.5$.

By definition, $\lambda(1,1) = 0.1$, $\lambda(1,2) = 0.2 + 0.3 = 0.5$, $\lambda(2,1) = 0.1$, $\lambda(2,2) = 0.1$, $\mu(1) = 0.7$, and $\mu(2) = 0.5$. Then $\rho(1) = 0.1/0.7 + 0.5/0.5 > 1$ and $\rho(2) = 0.1/0.7 + 0.1/0.5 < 1$. On the other hand, for $\boldsymbol{b} = (1,1)^{\mathrm{T}}$, $\max_{1\leqslant k\leqslant 2}\{((\Lambda - M)\boldsymbol{b})_k\} = -0.1$. Thus, this Markov chain is positive recurrent.

In Example 4.1, the Markov chain can go from node $J = 1$ to node $J = 2$ in one transition. This type of jump, from one branch of a tree to another without passing their parent node, makes the Markov chain more complicated. One would think that this is the reason why $\max_{1\leqslant k\leqslant K}\{\rho(k)\} < 1$ is not necessary for positive recurrence. The following example shows that $\max_{1\leqslant k\leqslant K}\{\rho(k)\} < 1$ is not necessary for positive recurrence when such transitions do not exist.

Example 4.2. Consider a Markov chain of M/G/1 type with a tree structure and $K=2$. Transition probabilities of this Markov chain are given as

$$a_0(1,1)=0.1,\quad a_0(1,2)=0.1,$$
$$a_0(2,1)=0.3,\quad a_0(2,2)=0.3,$$
$$a_2(1)=0.8,\quad a_2(2)=0.4.$$

Then $\lambda(1,1)=0.1$, $\lambda(1,2)=0.1$, $\lambda(2,1)=0.3$, $\lambda(2,2)=0.3$, $\mu(1)=0.8$, and $\mu(2)=0.4$. For this example, $\rho(1)=0.1/0.8+0.1/0.4<1$ and $\rho(2)=0.3/0.8+0.3/0.4=9/8>1$. On the other hand, for $\boldsymbol{b}=(1,3.1)^{\mathrm{T}}$, $\max_{1\leqslant k\leqslant 2}\{((\Lambda-M)\boldsymbol{b})_k\}=-0.1$. Thus, the Markov chain is positive recurrent.

The following example is presented to show that, sometimes, $\mathrm{sp}(P)$ can be found without solving any equation.

Corollary 4.2. *For the Markov chain defined in Section* 2, *assume that* $\lambda(k,j)=\lambda(j), 1\leqslant k,j\leqslant K$. *That is, the arrival rates are independent of the current node. Then* $\mathrm{sp}(P)=\rho(1)=\cdots=\rho(K)$ *and the classification of the Markov chain is determined by* $\rho(1)$.

Proof. In this case, it is clear that $\rho(1)=\cdots=\rho(K)$. By Theorem 3.2, we only have to prove that $\rho(1)$ is the Perron–Frobenius eigenvalue of matrix P. Let $b_k=1/\mu(k)$, $1\leqslant k\leqslant K$. It is easy to see that $(\Lambda-M)\boldsymbol{b}=\rho(1)\boldsymbol{e}-M\boldsymbol{b}$. Therefore, $\mathrm{sp}(P)=\rho(1)$. This completes the proof. □

Finally in this section, we look at the measure $\rho^*=\lambda(1,1)/\mu(1)+\cdots+\lambda(K,K)/\mu(K)$. This is usually the traffic intensity in queueing theory when multiple types of customers are present. In general, it is clear that ρ^* is not enough to classify the Markov chain, but it is enough for some special cases. As in the special case defined in Corollary 4.2, $\mathrm{sp}(P)=\rho^*$.

Another special case of interest is when all the service rates are the same, i.e., $\mu(1)=\cdots=\mu(K)$. Then $\rho^*=[\lambda(1,1)+\cdots+\lambda(K,K)]/\mu(1)$. One question is whether or not $\mathrm{sp}(P)=\rho^*$. This is not true in general. Here is a counterexample.

Example 4.3. Consider a Markov chain of M/G/1 type with a tree structure and $K=2$. Transition probabilities of this Markov chain are given as

$$a_0(1,1)=0.3,\quad a_0(1,2)=0.1,$$
$$a_0(2,1)=0.1,\quad a_0(2,2)=0.3,$$
$$a_2(1)=0.6,\ a_2(2)=0.6.$$

Then $\lambda(1,1)=0.3$, $\lambda(1,2)=0.1$, $\lambda(2,1)=0.1$, $\lambda(2,2)=0.3$, $\mu(1)=\mu(2)=0.6$. In this example, $\rho^*=0.3/0.6+0.3/0.6=1$. On the other hand, for $\boldsymbol{b}=(1,1)^{\mathrm{T}}$, $\max_{1\leqslant k\leqslant 2}\{((\Lambda-M)\boldsymbol{b})_k\}=-0.2$, which implies $\mathrm{sp}(P)<1$. Thus, the Markov chain is positive recurrent.

5. Queueing examples

In this section, the classification problem of a number of queueing systems is studied by using the results obtained in Sections 3 and 4. The key is to find the Perron–Frobenius eigenvalue of matrix P for the queueing systems. We begin with a simple discrete time queueing model.

Example 5.1. (*The* GEO[K]/GEO[K]/1/LCFS *preemptive resume queue*). Consider a discrete time single-server queueing system with a marked geometric arrival process (the discrete time counterpart of the marked Poisson process He and Neuts [7] and Neuts [13]) and geometric service times. All customers are served on a last-come-first-served (LCFS) preemptive resume basis. When a customer arrives (at the end of a period), if the current service is not completed, it pushes the current customer in the server out and starts its own service. When a customer in queue reenters the server, its service time resumes at the point it was pushed out.

Let p_k be the probability that a customer of type k arrives by the end of a period, $1\leqslant k\leqslant K$, and p_0 is the probability there is no customer arriving in a period. Then $\{p_0,p_1,\ldots,p_K\}$ describes the arrival process with $1=p_0+p_1+\cdots+p_K$. If a type k customer is in service at the beginning of a period, its service will be completed by the end of the period with probability q_k. Thus, the service times have geometric distributions and $\{q_1,\ldots,q_K\}$ describes the service processes.

Let X_n be the queue string at the beginning of period n, including the customer in service. Clearly $\{X_n,\ n\geqslant 0\}$ is a Markov chain of M/G/1 type with a tree structure. The transition probabilities of this

Markov chain are given as (after the decomposition defined by Eq. (3.12)): $a_0(k,j)=(1-q_k)p_j$, $a_1(k,j)=q_k p_j$, $a_1(k)=(1-q_k)p_0$, and $a_2(k)=q_k p_0$. Then the arrival rates and service rates are given as $\lambda(k,j)=p_j$ and $\mu(k)=q_k$. It is easy to see that $\mathrm{sp}(P)=\rho=p_1/q_1+\cdots+p_K/q_K$. Therefore, Markov chain $\{X_n,\ n\geqslant 0\}$ of the queueing system is positive recurrent if $\rho<1$, null recurrent if $\rho=1$, or transient if $\rho>1$.

Note. It is easy to see that $\rho=p_1/q_1+\cdots+p_K/q_K$ provides information for a complete classification of the GEO[K]/GEO[K]/1 queue with a FCFS, LCFS nonpreemptive, LCFS preemptive resume, or LCFS preemptive repeat service discipline.

Example 5.2 (*The* BGEO[K]/G[K]/1/LCFS *nonpreemptive queue*)**.** Consider a discrete time single-server queueing system with a marked batch geometric arrival process and general service times. All customers are served on a LCFS nonpreemptive basis. That is: the service of any customer will not be interrupted until it is completed.

Let p_J be the probability that a batch of type J arrives by the end of a period for $J\in\aleph$ and $J>0$, and p_0 the probability that there is no customer arriving in a period. Customers join the queue according to their order in the batch. Then $\{p_J,\ J\in\aleph\}$ describes the arrival process. The distribution of service time $S(k)$ of a type k customer is given by $\{q_k(n),\ n\geqslant 1\}$, i.e., $\boldsymbol{P}\{S(k)=n\}=q_k(n)$. The mean number $\boldsymbol{E}J(k)$ of type k customers who arrive during a unit period of time and the mean service time $\boldsymbol{E}S(k)$ of a type k customer are given as

$$\boldsymbol{E}J(k)=\sum_{H\in\aleph,\,H>0} p_H N(H,k) \quad \text{and}$$

$$\boldsymbol{E}S(k)=\sum_{n=1}^{\infty} n q_k(n). \tag{5.1}$$

We observe the queueing system at customer departure epochs and the epochs when arriving customers find an empty queue. Let X_n be the queue string right after the nth departure or empty queue epoch. $\{X_n,\ n\geqslant 0\}$ is an embedded Markov chain of M/G/1 type with a tree structure. Let $N(k)$ be the number of batches that arrived during the service period of a type k customer. Let $\xi(k)$ be the string consisting of all customers who arrived during the service period of a type k customer, i.e., $\xi(k)=J_1+\cdots+J_{N(k)}$, where J_i is the string of customers who arrived in the ith period. The distribution of $\xi(k)$ is given as, for $H\in\aleph$,

$$\boldsymbol{P}\{\xi(k)=H\} = \sum_{n=1}^{\infty} q_k(n)\left(\sum_{\{J_1,\dots,J_n\}\,:\,J_1+\cdots+J_n=H} p_{J_1}\cdots p_{J_n}\right). \tag{5.2}$$

The transition of Markov chain $\{X_n\}$ is given by $X_{n+1}=(J+k)-k+\xi(k)=J+\xi(k)$, given that $X_n=J+k$, and X_{n+1} is an arbitrary batch when $X_n=0$. By definition, it is clear that $\mu(k)=1$ and the arrival rates are obtained as (after the decomposition defined by Eq. (3.12))

$$\begin{aligned}
\lambda(k,j) &= \sum_{n=1}^{\infty} n\left(\sum_{H\,:\,N(H,j)=n} \boldsymbol{P}\{\xi(k)=H\}\right)\\
&= \sum_{n=1}^{\infty} n\left[\sum_{N(H,j)=n}\ \sum_{t=1}^{\infty} q_k(t)\right.\\
&\qquad \left.\times\left(\sum_{\{J_1,\dots,J_t\}\,:\,J_1+\cdots+J_t=H} p_{J_1}\cdots p_{J_t}\right)\right]\\
&= \sum_{t=1}^{\infty} q_k(t)\left[\sum_{n=1}^{\infty}\ \sum_{H\,:\,N(H,j)=n}\right.\\
&\qquad \left.\times\, n\left(\sum_{\{J_1,\dots,J_t\}\,:\,J_1+\cdots+J_t=H} p_{J_1}\cdots p_{J_t}\right)\right]\\
&= \sum_{t=1}^{\infty} q_k(t)\left[\sum_{n=1}^{\infty}\ \sum_{H\,:\,N(H,j)=n}\right.\\
&\qquad \times\left(\sum_{\{J_1,\dots,J_t\}\,:\,J_1+\cdots+J_t=H}\right.\\
&\qquad \left.\left.\times\,[N(J_1,j)+\cdots+N(J_t,j)]\,p_{J_1}\cdots p_{J_t}\right)\right]\\
&= \sum_{t=1}^{\infty} q_k(t)t\boldsymbol{E}J(j)=\boldsymbol{E}S(k)\boldsymbol{E}J(j).
\end{aligned} \tag{5.3}$$

Eq. (5.3) is intuitive since the average total number of type j customers who arrived during the service time of a type k customer is the average number of type j customers who arrive in a unit period of time multiplying the mean service time of the type k customer. Then matrix P of the corresponding Markov chain is given as

$$P = M^{-1}\Lambda = \begin{pmatrix} \boldsymbol{E}S(1) \\ \vdots \\ \boldsymbol{E}S(K) \end{pmatrix} (\boldsymbol{E}J(1) \quad \cdots \quad \boldsymbol{E}J(K)). \tag{5.4}$$

Because of the special structure, the Perron–Frobenius eigenvalue of matrix P is given as $\mathrm{sp}(P) = \rho = \sum_{k=1}^{K} \boldsymbol{E}S(k)\boldsymbol{E}J(k)$ (the classical traffic intensity of the queueing system), which determines the classification of the queueing system.

Note. It is easy to see that the classification of the BGEO[K]/G[K]/1 queue with a FCFS He [4], LCFS preemptive resume, or other work conserving service discipline is the same as that of Example 5.2.

Note. Let us consider a BGEO[K]/G[K]/1 queue with $K=3$, arrival rates $\boldsymbol{E}J(1)=1/12$, $\boldsymbol{E}J(3)=1/6$, $\boldsymbol{E}J(3)=1/4$, and mean service times $\boldsymbol{E}S(1)=4$, $\boldsymbol{E}S(2)=2$, and $\boldsymbol{E}S(3)=1$. Then $\rho = \boldsymbol{E}S(1)\boldsymbol{E}J(1) + \boldsymbol{E}S(2)\boldsymbol{E}J(2) + \boldsymbol{E}S(3)\boldsymbol{E}J(3) = 11/12 < 1$, which implies that the queueing system is positive recurrent. Consider a BGEO[1]/G[1]/1 queue with arrival rate $\lambda = \boldsymbol{E}J(1) + \boldsymbol{E}J(2) + \boldsymbol{E}J(3) = 0.5$ and mean service time $1/\mu = w_1\boldsymbol{E}S(1) + w_2\boldsymbol{E}S(2) + w_3\boldsymbol{E}S(1)$, where $w_1, w_2, w_3 \geqslant 0$ and $w_1 + w_2 + w_3 = 1$. If we choose $w_1 = w_2 = w_3 = 1/3$, $\lambda/\mu = 7/6 > 1$. If we choose $w_1 = 0$, $w_2 = 0$, and $w_3 = 1$, $\lambda/\mu = 0.5$. If we choose $w_1 = 1/6$, $w_2 = 1/3$, $w_3 = 1/2$, $\lambda/\mu = 11/12 = \rho$. Thus, the traffic intensity of the queueing system with a single type of customer depends on the set of weights used in estimating the service time of an arbitrary customer. In general, λ/μ can be far away from ρ of the original queueing system. The two are equal (i.e., $\lambda/\mu = \rho$) when weights are chosen as $w_k = \boldsymbol{E}J(k)[\boldsymbol{E}J(1) + \cdots + \boldsymbol{E}J(K)]^{-1}$, $1 \leqslant k \leqslant K$.

So far in this section, we have shown that the classification of some queueing systems is determined completely by the classical traffic intensity ρ. The reason is that these queueing systems are work conserving. In the next example, the classification problem of a queueing system which is not work conserving is discussed.

Example 5.3 (*The* BGEO[K]/G[K]/1/LCFS *preemptive repeat queue*)**.** Consider a discrete time queueing system with a marked batch geometric arrival process, general service times, and a LCFS preemptive repeat service discipline. When a customer reenters the server, its service time has the same distribution as that of a new customer. Parameters defined in Example 5.2 are used.

Let X_n be the queue string at the nth service completion or arrival epoch. $\{X_n,\ n \geqslant 0\}$ is a Markov chain because of the preemptive repeat service discipline. The transition probabilities of this Markov chain are given as, for $J > 0$, (after the decomposition defined by Eq. (3.12))

$$a_0(k,J) = \sum_{n=1}^{\infty} p_0^{n-1} p_J \left(\sum_{t=n+1}^{\infty} q_k(t) \right),$$

$$a_1(k,J) = \sum_{n=1}^{\infty} p_0^{n-1} p_J q_k(n), \tag{5.5}$$

$$a_2(k) = \sum_{n=1}^{\infty} p_0^{n} q_k(n).$$

Denote by $q_k^*(z) = \sum_{n=1}^{\infty} z^n q_k(n)$. Then we obtain

$$\begin{aligned}\mu(k) &= \sum_J a_1(k,J) + a_2(k) \\ &= \sum_J \sum_{n=1}^{\infty} p_0^{n-1} p_J q_k(n) + \sum_{n=1}^{\infty} p_0^{n} q_k(n) \\ &= \sum_{n=1}^{\infty} p_0^{n-1} q_k(n) = q_k^*(p_0)/p_0, \end{aligned} \tag{5.6}$$

$$\begin{aligned}\lambda(k,j) &= \sum_{n=1}^{\infty} p_0^{n-1} q_k(n) \boldsymbol{E}J(j) \\ &\quad + \sum_{t=2}^{\infty} q_k(t) \left(\sum_{n=1}^{t-1} p_0^{n-1} \boldsymbol{E}J(j) \right) \\ &= \frac{(1 - q_k^*(p_0))}{(1 - p_0)} \boldsymbol{E}J(j), \end{aligned} \tag{5.7}$$

then

$$P = M^{-1}\Lambda = \frac{p_0}{1-p_0}\begin{pmatrix} \dfrac{1-q_1^*(p_0)}{q_1^*(p_0)} \\ \vdots \\ \dfrac{1-q_K^*(p_0)}{q_K^*(p_0)} \end{pmatrix} \times (\boldsymbol{E}J(1) \quad \cdots \quad \boldsymbol{E}J(K)). \tag{5.8}$$

Therefore, the Perron–Frobenius eigenvalue of matrix P is given as

$$\mathrm{sp}(P) = \begin{cases} \dfrac{p_0}{(1-p_0)} \sum_{k=1}^{K} \dfrac{(1-q_k^*(p_0))}{q_k^*(p_0)} \boldsymbol{E}J(k), & 0 < p_0 < 1, \\ \sum_{k=1}^{K} \dfrac{\boldsymbol{E}J(k)}{q_k(1)}, & p_0 = 0. \end{cases} \tag{5.9}$$

According to Theorem 3.2, expression (5.9) provides complete information for the classification of the queueing system of interest.

Next, we consider continuous time counterparts of the two queueing systems considered in Examples 5.2 and 5.3.

Example 5.4 (*The* BM[K]/G[K]/1/LCFS *nonpreemptive queue*)**.** Consider a queueing system with a marked batch Poisson arrival process and general service times. Let p_J be the arrival rate of a batch of type J for $J \in \aleph$ and $J > 0$. Then $\{p_J,\ J \in \aleph,\ J > 0\}$ describes the Poisson arrival process. Note that p_J can be any nonnegative real number in this and the next examples. Let $p_0 = \sum_{J>0} p_J$. Then p_0 is the parameter of the exponential distribution of interarrival times. The service time of a type k customer is given by $F_k(t)$ with Laplace Stieltjes transform $f_k^*(s)$. Customers are served on a LCFS and nonpreemptive basis. The mean arrival rate of type k customers and the mean service time of a type k customer are given as

$$\boldsymbol{E}J(k) = \sum_{n=1}^{\infty} n \left(\sum_{H:N(H,k)=n} p_H \right), \qquad \boldsymbol{E}S(k) = \int_0^{\infty} tF_k(\mathrm{d}t), \tag{5.10}$$

respectively. Let X_n be the queue string right after the nth departure epoch. Then $\{X_n,\ n \geqslant 0\}$ is a Markov chain of M/G/1 type with a tree structure. The distribution of $\xi(k)$ is given as, for $H \in \aleph$,

$$\boldsymbol{P}\{\xi(k) = H\} = \int_0^{\infty} F_k(\mathrm{d}t) \sum_{n=0}^{\infty} \frac{\exp\{-p_0 t\}(p_0 t)^n}{n!} \times \left(\sum_{\{J_1,\ldots,J_n\}:J_1+\cdots+J_n=H} \frac{p_{J_1}}{p_0} \cdots \frac{p_{J_n}}{p_0} \right). \tag{5.11}$$

Then the transition of Markov chain $\{X_n,\ n \geqslant 0\}$ is given by $X_{n+1} = (J+k) - k + \xi(k) = J + \xi(k)$ when $X_n = J+k$. By definition, it is clear that $\mu(k)=1$ and the arrival rates are obtained as (after the decomposition define by Eq. (3.12))

$$\begin{aligned}
\lambda(k,j) &= \sum_{n=1}^{\infty} n \left(\sum_{H:N(H,j)=n} \boldsymbol{P}\{\xi(k) = H\} \right) \\
&= \sum_{n=1}^{\infty} n \sum_{H:N(H,j)=n} \int_0^{\infty} F_k(\mathrm{d}t) \times \sum_{m=0}^{\infty} \frac{\exp\{-p_0 t\}(p_0 t)^m}{m!} \times \left(\sum_{\{J_1,\ldots,J_m\}:J_1+\cdots+J_m=H} \frac{p_{J_1}}{p_0} \cdots \frac{p_{J_m}}{p_0} \right) \\
&= \int_0^{\infty} F_k(\mathrm{d}t) \sum_{m=1}^{\infty} \frac{\exp\{-p_0 t\}(p_0 t)^m}{m!} \sum_{n=1}^{\infty} n \times \sum_{H:N(H,j)=n} \left(\sum_{\{J_1,\ldots,J_m\}:J_1+\cdots+J_m=H} \frac{p_{J_1}}{p_0} \cdots \frac{p_{J_m}}{p_0} \right) \\
&= \int_0^{\infty} F_k(\mathrm{d}t) \sum_{m=1}^{\infty} \frac{\exp\{-p_0 t\}(p_0 t)^m}{m!} \sum_{n=1}^{\infty} n \times \sum_{H:N(H,j)=n} \left(\sum_{\{J_1,\ldots,J_m\}:J_1+\cdots+J_m+H} [N(J_1,j) + \cdots + N(J_m,j)] \frac{p_{J_1}}{p_0} \cdots \frac{p_{J_m}}{p_0} \right)
\end{aligned}$$

$$= \int_0^\infty F_k(\mathrm{d}t) \sum_{m=1}^{\infty} \frac{\exp\{-p_0 t\}(p_0 t)^m}{m!} \times \left(\frac{m\boldsymbol{E}J(j)}{p_0}\right) = \boldsymbol{E}S(k)\boldsymbol{E}J(j). \quad (5.12)$$

It is clear that $\mathrm{sp}(P)=\rho=\sum_{k=1}^{K} \boldsymbol{E}S(k)\boldsymbol{E}J(k)$ determines the classification of the queueing system of interest. Again, this result applies to any BM[K]/G[K]/1 queue with a work-conserving service discipline.

Example 5.5 (*The* BM[K]/G[K]/1/LCFS *preemptive repeat queue*)**.** Consider the queueing system defined in Example 5.4 when customers are served on a LCFS preemptive repeat basis. Let X_n be the queue string at the nth service completion or arrival epoch. Then $\{X_n,\ n\geqslant 0\}$ is a Markov chain because of the preemptive repeat service discipline. The transition probabilities of this Markov chain are given as

$$a_0(k,J) = \int_0^\infty (1-\exp\{-p_0 t\})F_k(\mathrm{d}t)\frac{p_J}{p_0}, \qquad a_1(k)=0, \quad a_2(k) = \int_0^\infty \exp\{-p_0 t\}F_k(\mathrm{d}t). \quad (5.13)$$

Then we obtain

$$\mu(k) = a_2(k) = f_k^*(p_0), \quad (5.14)$$

$$\lambda(k,j) = \int_0^\infty (1-\exp\{-p_0 t\})F_k(\mathrm{d}t)\frac{\boldsymbol{E}J(j)}{p_0} = \frac{(1-f_k^*(p_0))}{p_0}\boldsymbol{E}J(j). \quad (5.15)$$

Therefore, the Perron–Frobenius eigenvalue of matrix P is given as

$$\mathrm{sp}(P) = \sum_{k=1}^{K} \frac{(1-f_k^*(p_0))}{f_k^*(p_0)} \frac{\boldsymbol{E}J(k)}{p_0}, \quad 0 < p_0 < \infty. \quad (5.16)$$

According to Theorem 3.2, the classification of the Markov chain or the queueing system is determined completely by the expression in Eq. (5.16).

6. Continuous time Markov chain of M/G/1 type with a tree structure

In this section, the continuous time Markov chains (Markov processes) of M/G/1 type with a tree structure are discussed briefly. Results obtained in Sections 3–5 hold for the continuous time case. Some of them are restated for the continuous time case in this section.

A Markov process $\{X(t),\ t\geqslant 0\}$ is of M/G/1 type with a tree structure if it is defined on a K-ary tree and its transition rates are described as follows: the transition from node $J+k$ to

(1) $J+k+H$ with rate $a_0(k,H)$ when $H\in\aleph$ and $H>0$;
(2) $J+H$ with rate $a_1(k,H)$ when $H=j_1\cdots j_{|H|}\in\aleph$ and $0<j_1\neq k$;
(3) J with rate $a(J+k,J)=a_2(k)$;
(4) H with rate $a_0(0,H)$ when $J=0$, $H\in\aleph$ and $H>0$.

The sojourn time in node $J+k$ has an exponential distribution with parameter $-[\sum_H (a_0(k,H)+a_1(k,H))+a_2(k)]$ and $-\sum_H a_0(0,H)$ for the root node.

$\{\lambda(k,j), \mu(k), \Lambda, M, P, \mathrm{sp}(P)\}$ are the same as those defined in Section 2. We assume that Markov process $X(t)$ is irreducible, $\{\lambda(k,j),\ \mu(k),\ 0\leqslant k\leqslant K,\ 1\leqslant j\leqslant K\}$ are finite, and matrix P is irreducible.

Theorem 6.1. *For the Markov process of M/G/1 type with a tree structure $\{X(t),\ t\geqslant 0\}$ defined above, it is*

(1) *positive recurrent if and only if* $\mathrm{sp}(P)<1$;
(2) *null recurrent if and only if* $\mathrm{sp}(P)=1$;
(3) *transient if and only* $\mathrm{sp}(P)>1$.

Proof. The proof is similar to that of Theorem 3.2. Details are omitted.

Example 6.1 (*The* M[K]/M[K]/1 *queue*)**.** Consider a queueing system with a marked Poisson arrival process and exponential service times. Let p_k be the arrival rate of type k customers, $1\leqslant k\leqslant K$. Then $\{p_0, p_1, \ldots, p_K\}$ describes the arrival process with $p_0 = p_1+\cdots+p_K$. If a type k customer is in service, its service time has an exponential distribution with parameter q_k. Suppose that the LCFS preemptive resume service discipline is applied.

Let $X(t)$ be the queue string at time t. Then the infinitesimal generator of Markov process $\{X(t),\ t\geqslant 0\}$ are given as: $a_0(k,j)=p_j$, $a_1(k)=-q_k-p_0$, and $a_2(k)=q_k$. The total arrival rates and service rates are given as $\lambda(k,j)=p_j$ and $\mu(k)=q_k$. It is easy to see that $\mathrm{sp}(P)=\rho=p_1/q_1+\cdots+p_K/q_K$ determines the classification of the Markov process or the queueing system.

Note. It is clear that the result can be generalized to a batch arrival case. It is also easy to see that the result can be extended to the multiple server case, i.e., the BM[K]/M[K]/s queue with a FCFS, LCFS nonpreemptive, or LCFS preemptive resume service discipline is classified by ρ.

Note. It is easy to see that the decomposition technique can be used in the continuous time case.

Acknowledgements

The author would like to thank Dr. Bhaskar Sengupta for his valuable comments and suggestions.

References

[1] G. Fayolle, V.A. Malyshev, M.V. Menshikov, Topics in the Constructive Theory of Countable Markov Chains, Cambridge University Press, Cambridge, 1995.
[2] A.S. Gajrat, V.A. Malyshev, M.V. Menshikov, K.D. Pelih, Classification of Markov chains describing the evolution of a string of characters, Uspehi Matematicheskih Nauk 50 (2) (1995) 5–24.
[3] F.R. Gantmacher, The Theory of Matrices, Chelsea, New York, 1959.
[4] Q.-M. He, Queues with marked customers, Adv. Appl. Proab. 28 (1996) 567–587.
[5] Q.-M. He, Classification of Markov processes of matrix M/G/1 type with a tree structure and its applications to queueing models, submitted for publication.
[6] Q.-M. He, A.S. Alfa, The MMAP[K]/PH[K]/1 queues with a last-come-first-served preemptive service discipline, Queueing Systems 29 (1998) 269–291.
[7] Q.-M. He, M.F. Neuts, Markov arrival processes with marked transitions, Stochastic Process Appl. 74/1 (1998) 37–52.
[8] J.G. Kemeny, J.L. Snell, A.W. Knapp, Denumerable Markov Chains, D. Van Nostrand Company, Inc., Princeton, NJ, 1966.
[9] P.R. Kumar, S.P. Meyn, Duality and linear programs for stability and performance analysis of queueing networks and scheduling policies, IEEE Trans. Automat. Control 41 (1) (1996) 4–17.
[10] V.A. Malyshev, Interacting strings of symbols, Russian Math. Surveys 52 (2) (1997) 59–86.
[11] V.A. Malyshev, Stochastic evolution via graph grammars, INRIA Research Report #3380, 1998.
[12] S.P. Meyn, R. Tweedie, Markov Chains and Stochastic Stability, Springer, Berlin, 1993.
[13] M.F. Neuts, A versatile Markovian point process, J. Appl. Probab. 16 (1979) 764–779.
[14] M.F. Neuts, Matrix-Geometric Solutions in Stochastic Models: An algorithmic Approach, The Johns Hopkins University Press, Baltimore, 1981.
[15] M.F. Neuts, Structured Stochastic Matrices of M/G/1 type and Their Applications, Marcel Dekker, New York, 1989.
[16] T. Takine, B. Sengupta, R.W. Yeung, A generalization of the matrix M/G/1 paradigm for Markov chains with a tree structure, Stochastic Models 11 (1995) 411–421.
[17] R.W. Yeung, B. Sengupta, Matrix product-form solutions for Markov chains with a tree structure, Adv. Appl. Probab. 26 (4) (1994) 965–987.