
Volume 36, Issue 6, November 2008
ISSN 0167-6377
ELSEVIER
Operations Research Letters
EDITOR-IN-CHIEF
Jan Karel Lenstra
AREA EDITORS
approximation and heuristics
Gerhard Woeginger
book reviews
Frits Spieksma
continuous optimization
Patrice Marcotte
financial engineering
Steven Kou
game theory
Nimrod Megiddo
graphs and networks
Bruce Shepherd
inventory, reliability and control
Sridhar Seshadri
linear and stochastic optimization
Rüdiger Schultz
location, routing and supply chains
Ton de Kok
mixed integer optimization
Christoph Helmberg and
Alexander Martin
sequencing and scheduling
Steef van de Velde
simulation
Sigrún Andradóttir
stochastic networks and queues
Sem Borst and Tetsuya Takine
www.elsevier.com/locate/orl

Operations Research Letters 36 (2008) 705–709

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Stability of a priority queueing system with customer transfers

Jingui Xie [a], Qi-Ming He [b], Xiaobo Zhao [a,*]

[a] Department of Industrial Engineering, Tsinghua University, Beijing 100084, PR China
[b] Department of Industrial Engineering, Dalhousie University, Halifax, N.S., Canada B3J 2X4

ARTICLE INFO

Article history:
Received 17 January 2008
Accepted 19 June 2008
Available online 18 July 2008

Keywords:
Priority queue
Markov chain
Mean-drift method
Stability
Ergodicity

ABSTRACT

This paper is concerned with the stability of a preemptive priority queueing system with customer transfers. Conditions for the queueing system to be stable/unstable are found. An interesting result is that the stability/instability conditions are independent of the service rates of lower priority customers and the transfer rates.

1. Introduction

The queueing model of interest has N types of customers: type 1, type 2, ..., and type N customers. The priority increases from type 1 customers to type N customers, with type 1 customers having the lowest priority and type N customers the highest priority. This paper finds simple conditions for the stability/instability of the queueing model.

The queueing model of interest can find applications in the design of emergency departments in healthcare systems and custom inspection systems. For example, in a hospital emergency department, patients are categorized into critical and non-critical groups. A patient in critical condition will be attended by a doctor, if one is available, as soon as the patient arrives. The status of a patient in non-critical condition may deteriorate and become critical. For a custom inspection system, perishable products such as food require immediate attention. Other products may be expedited while waiting and need to be inspected at the earliest available time. In such systems, items have different service priorities. The service priority of an item may increase while waiting.

Queueing systems with customer priorities and queueing systems with customer transfers have wide applications in manufacturing, computer networks, telecommunication systems, and vehicle traffic control. The study of such queueing systems is extensive. Existing works address issues related to system stability, optimal scheduling, routing, and performance analysis [1,9,10,12,14–17]. For example, some of the existing works focus on system stability conditions, some on the stationary analysis of the queue length(s) and waiting times, and some on customer transfer strategies. The queueing model of interest is also related to, but not included in, stochastic transfer networks [2]. In addition, the literatures focus on a product-form solution, rather than stability conditions.

In the priority queueing system of interest, a customer transfer scheme is given. The model is different from those in the existing literature. As a consequence, the stability/instability conditions of the system are different from those of the existing models. An interesting result is that the stability/instability conditions depend only on the ratio of the sum of the arrival rates and the service rate of customers of the highest priority, i.e., the stability/instability conditions are independent of the service rates of lower priority customers and the transfer rates. That result can be useful in the design of such queueing systems.

The mean-drift method is the main mathematical tools used in this paper [5–8,11]. This method has been used in the past in the study of classification of Markov processes and queueing models. One of the keys in using this method is the construction of the Lyapunov (test) functions. In this paper, several Lyapunov functions are introduced and they lead to the findings of simple conditions for system stability and instability.

The remainder of the paper is organized as follows. The queueing model of interest is introduced in Section 2. The main results – stability and instability conditions – are presented. Section 3 gives proofs of the main results.

* Corresponding author. Tel.: +86 10 6278 4898; fax: +86 10 6279 4399.
E-mail address: xbzhao@tsinghua.edu.cn (X. Zhao).

0167-6377/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.orl.2008.06.007

2. Queueing model and main results

The queueing model of interest consists of s identical servers serving N types of customers: type 1, type 2, …, and type N customers. Type $1, 2, \ldots,$ and N customers form queue $1, 2, \ldots,$ and N, respectively. Type N customers have the highest service priority, type $N-1$ the second highest service priority, …, and type 1 the lowest service priority. When a server is available, it chooses a customer from the non-empty queue of the highest priority and begins to serve it. If some servers are serving type j customers when a type k customer arrives, for $j < k$, there is no idle server, and type j customers are the lowest priority customers in service, then one of the type j customers in service is pushed back to queue j and the server begins to serve the type k customer. The type j customer will resume or repeat its service if a server is available to serve type j customers.

Type $1, 2, \ldots,$ and N customers arrive according to independent Poisson processes with parameters $\lambda_1, \lambda_2, \ldots,$ and λ_N, respectively. The service times of type $1, 2, \ldots,$ and N customers are exponentially distributed with parameters $\mu_1, \mu_2, \ldots,$ and μ_N, respectively. The arrival processes and service times are independent. Since the service time of a type j customer is exponentially distributed, it does not make a difference to assume that its interrupted service, if it occurs, will be repeated or resumed. For the same reason, if a server is available to serve type j customers, it does not matter (to system stability/instability) which waiting type j customer enters the server to receive service.

While waiting in queue, a type j customer may change to a type $j+1$ customer after an exponential time with parameter $\lambda_{T,j}$, for $1 \le j \le N-1$. Since the time before transfer is exponentially distributed, it does not make a difference to assume that the clock until transfer is reset or continued, if a type j customer's service is interrupted. The times until transfers for individual customers are independent of each other, and are independent of the arrival and service processes. Note that a customer in service does not change its type.

Define $q_j(t)$ the number of type j customers in queue j at time t, which includes the customers in service (if there are customers in service), $j = 1, 2, \ldots, N$. If all system parameters are positive, it is easy to see that $\{(q_1(t), q_2(t), \ldots, q_N(t)), t \ge 0\}$ is an irreducible continuous time Markov chain (CTMC) with a state space $\{(q_1, q_2, \ldots, q_N), q_1 \ge 0, \ldots, q_N \ge 0\}$. Denote by $Q = (Q_{(q_1,q_2,\ldots,q_N),(y_1,y_2,\ldots,y_N)})$ the infinitesimal generator of the Markov chain. We have, for $(q_1, q_2, \ldots, q_N) \neq (y_1, y_2, \ldots, y_N)$,

$$Q_{(q_1,q_2,\ldots,q_N),(y_1,y_2,\ldots,y_N)} = \begin{cases} \lambda_j, \quad \text{if } y_j = q_j + 1, y_i = q_i, i \neq j, 1 \le j \le N; \\ \min\{s, q_N\}\mu_N, \quad \text{if } y_j = q_j,\ 1 \le j \le N-1, \\ \quad y_N = q_N - 1 \ge 0; \\ \min\left\{s - \displaystyle\sum_{k=j+1}^{N} q_k, q_j\right\}\mu_j, \quad \text{if } y_j = q_j - 1 \ge 0,\ y_i = q_i, \\ \quad i \neq j,\ \displaystyle\sum_{k=j+1}^{N} q_k < s, 1 \le j \le N; \\ q_j\lambda_{T,j}, \quad \text{if } y_j = q_j - 1 \ge 0,\ y_{j+1} = q_{j+1} + 1, y_i = q_i, i \neq j, \\ \quad j+1,\ \displaystyle\sum_{k=j+1}^{N} q_k \ge s, 1 \le j \le N-1; \\ \max\left\{\displaystyle\sum_{k=j}^{N} q_k - s, 0\right\}\lambda_{T,j}, \quad \text{if } y_j = q_j - 1 \ge 0, \\ \quad y_{j+1} = q_{j+1} + 1, y_i = q_i, i \neq j, j+1, \\ \displaystyle\sum_{k=j+1}^{N} q_k < s,\ 1 \le j \le N; \\ 0, \quad \text{otherwise}. \end{cases}$$

We say that the queueing system is stable if the Markov chain $\{(q_1(t), q_2(t), \ldots, q_N(t)), t \ge 0\}$ is ergodic (i.e., irreducible and positive recurrent). The Markov chain is called non-ergodic if it is not ergodic. The ergodicity of the Markov chain $\{(q_1(t), q_2(t), \ldots, q_N(t)), t \ge 0\}$ is characterized in the following theorem.

Theorem 1. *Assume that all system parameters* $\{\lambda_1, \ldots, \lambda_N, \mu_1, \ldots, \mu_N, \lambda_{T,1}, \ldots, \lambda_{T,N}, s\}$ *are positive and finite. The Markov chain* $\{(q_1(t), q_2(t), \ldots, q_N(t)), t \ge 0\}$ *is irreducible.*

(1.1) *The Markov chain is ergodic if* $\sum_{j=1}^{N} \lambda_j < s\mu_N$.

(1.2) *The Markov chain is non-ergodic if* $\sum_{j=1}^{N} \lambda_j > s\mu_N$.

(1.3) *The Markov chain is non-ergodic if* $\sum_{j=1}^{N} \lambda_j = s\mu_N$ *and* $\max_{1\le j\le N-1}\{\mu_j\} \le \mu_N$.

(1.4) *The Markov chain is recurrent if* $\sum_{j=1}^{N} \lambda_j = s\mu_N$ *and* $\min_{1\le j\le N-1}\{\mu_j\} \ge \mu_N$.

Part (1.1) and part (1.2) indicate that the ergodicity/non-ergodicity conditions of the Markov chain (or the stability/instability of the queueing system) are independent of the service rates of lower priority customers $\mu_1, \mu_2, \ldots,$ and μ_{N-1}. Therefore, system stability is determined by arrival rates (of all customers) and the service rate of type N customers.

Note 1: Although we assume a preemptive service discipline, all results hold for the system with a non-preemptive service discipline. The proofs are similar but more tedious, though.

3. Proof of Theorem 1

In this proof of Theorem 1, the mean-drift method [6] is utilized. Theorem 1.18 in Chen [3] for ergodicity of Markov chains, Theorem 1 in Choi and Kim [4] for non-ergodicity, and Part (ii) of Theorem 2.2 in Tweedie [13] for recurrence are applied.

Proof of (1.1). The proof of (1.1) includes three steps: (i) Construction of a Lyapunov function (test function); (ii) Calculation of the mean drift; and (iii) Application of Theorem 1.18 in [3].

We begin with the selection of the parameters in the Lyapunov function. Define

$$h(x) = x\sum_{j=1}^{N} \lambda_j + \frac{s\mu_N}{x} - \left(s\mu_N + \sum_{j=1}^{N} \lambda_j\right), \quad x > 0. \tag{1}$$

By routine calculations, it is easy to obtain $h(1) = h(x^*) = 0$, where $x^* = s\mu_N \left(\sum_{j=1}^{N} \lambda_j\right)^{-1} > 1$, and $h^{(1)}(1) = \lambda_1 + \cdots + \lambda_N - s\mu_N < 0$. Since the function $h(x)$ is convex for $x > 0$, we have $h(x) < 0$ for $1 < x < x^*$. Choose constants $\{a_1, a_2, \ldots, a_N\}$ satisfying the following conditions:

$$\begin{aligned} &(1) \quad 1 < a_N < a_{N-1} < \cdots < a_1 < x^*; \\ &(2) \quad h(a_N) + \sum_{j=1}^{N} \lambda_j (a_j - a_N) \equiv -\varepsilon < 0. \end{aligned} \tag{2}$$

Note that conditions in Eq. (2) can be satisfied since $h(a_N) < 0$.

For convenience, let $\mathbf{q} = (q_1, \ldots, q_N)$, $\mathbf{y} = (y_1, \ldots, y_N)$, and $\mathbf{e}(j)$ be the row vector with all elements being zero except that the jth element is one.

Now, we are ready to introduce the following Lyapunov function:

$$f(\mathbf{q}) = \begin{cases} \displaystyle\prod_{j=1}^{N} a_j^{q_j}, & q_j \ge 0,\ 1 \le j \le N; \\ 0, & \text{otherwise}. \end{cases}$$

It is easy to show that $f(\mathbf{q})$ is a compact function (see [3]), i.e., for any finite $d > 0$, the set $\{\mathbf{q} : f(\mathbf{q}) \le d, \mathbf{q} \ge 0\}$ is closed and bounded (if the variables $q_1, \ldots,$ and q_N take real values). Note that $\mathbf{q} \geq 0$ is equivalent to $q_1 \geq 0, \ldots,$ and $q_N \geq 0$. It is also easy to see that set $\{\mathbf{q} : f(\mathbf{q}) \leq d, \mathbf{q} \geq 0\}$ has a finite number of elements if variables $q_1, \ldots,$ and q_N take integer values and d is finite.

Next, we calculate the mean drift at a given state $\mathbf{q} = (q_1, \ldots, q_N)$, which is defined for the continuous time Markov chain as $\sum_{\mathbf{y}\geq 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y})$, where $Q_{\mathbf{q},\mathbf{y}} = Q_{(q_1,\ldots,q_N),(y_1,\ldots,y_N)}$. We need to consider the following three cases: (i) $q_N \geq s$, (ii) $q_N < s$ and $\sum_{j=1}^{N} q_j \geq s$, and (iii) $\sum_{j=1}^{N} q_j < s$.

If $q_N \geq s$, we have the following calculations:

$$\begin{aligned}
&\sum_{\mathbf{y}\geq 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) \\
&= s\mu_N f(\mathbf{q}-\mathbf{e}(N)) + \sum_{j=1}^{N} \lambda_j f(\mathbf{q}+\mathbf{e}(j)) + \sum_{j=1}^{N-1} q_j \lambda_{T,j} \\
&\quad \times f(\mathbf{q}+\mathbf{e}(j+1)-\mathbf{e}(j)) \\
&\quad - \left(s\mu_N + \sum_{j=1}^{N} \lambda_j + \sum_{j=1}^{N-1} q_j \lambda_{T,j} \right) f(\mathbf{q}) \\
&= f(\mathbf{q}) \left(\frac{s\mu_N}{a_N} + \sum_{j=1}^{N} \lambda_j a_j + \sum_{j=1}^{N-1} q_j \lambda_{T,j} \frac{a_{j+1}}{a_j} \right. \\
&\quad \left. - \left(s\mu_N + \sum_{j=1}^{N} \lambda_j + \sum_{j=1}^{N-1} q_j \lambda_{T,j} \right) \right) \\
&= f(\mathbf{q}) \left(\frac{s\mu_N}{a_N} + \sum_{j=1}^{N} \lambda_j a_j + \sum_{j=1}^{N-1} q_j \lambda_{T,j} \left(\frac{a_{j+1}}{a_j} - 1 \right) \right. \\
&\quad \left. - \left(s\mu_N + \sum_{j=1}^{N} \lambda_j \right) \right) \\
&\leq f(\mathbf{q}) \left(\frac{s\mu_N}{a_N} + \sum_{j=1}^{N} \lambda_j a_j - \left(s\mu_N + \sum_{j=1}^{N} \lambda_j \right) \right) \\
&= f(\mathbf{q}) \left(h(a_N) + \sum_{j=1}^{N-1} \lambda_j (a_j - a_N) \right).
\end{aligned} \tag{3}$$

Note that the fact $a_{j+1} < a_j$ is used for the inequality in Eq. (3).

If $q_N < s$ and $\sum_{j=1}^{N} q_j \geq s$, there exists k such that $\sum_{j=k+1}^{N} q_j < s$, $\sum_{j=k}^{N} q_j \geq s$, for $1 \leq k \leq N-1$. For this case, we have

$$\begin{aligned}
&\sum_{\mathbf{y}\geq 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) \\
&= \min\left\{ s - \sum_{j=k+1}^{N} q_j, q_k \right\} \mu_k f(\mathbf{q}-\mathbf{e}(k)) \\
&\quad + \sum_{j=k+1}^{N} q_j \mu_j f(\mathbf{q}-\mathbf{e}(j)) + \sum_{j=1}^{N} \lambda_j f(\mathbf{q}+\mathbf{e}(j)) \\
&\quad + \left(\sum_{j=k}^{N} q_j - s \right) \lambda_{T,k} f(\mathbf{q}+\mathbf{e}(k+1)-\mathbf{e}(k)) \\
&\quad + \sum_{j=1}^{k-1} q_j \lambda_{T,j} f(\mathbf{q}+\mathbf{e}(j+1)-\mathbf{e}(j)) \\
&\quad - \left(\min\left\{ s - \sum_{j=k+1}^{N} q_j, q_k \right\} \mu_k + \sum_{j=k+1}^{N} q_j \mu_j \right. \\
&\quad \left. + \sum_{j=1}^{N} \lambda_j + \left(\sum_{j=k}^{N} q_j - s \right) \lambda_{T,k} + \sum_{j=1}^{k-1} q_j \lambda_{T,j} \right) f(\mathbf{q}) \\
&= f(\mathbf{q}) \left(\left(s - \sum_{j=k+1}^{N} q_j \right) \frac{\mu_k}{a_k} + \sum_{j=k+1}^{N} \frac{q_j \mu_j}{a_j} + \sum_{j=1}^{N} \lambda_j a_j \right. \\
&\quad \left. + \left(\sum_{j=k}^{N} q_j - s \right) \frac{\lambda_{T,k} a_{k+1}}{a_k} + \sum_{j=1}^{k-1} \frac{q_j \lambda_{T,j} a_{j+1}}{a_j} \right) \\
&\quad - f(\mathbf{q}) \left(\left(s - \sum_{j=k+1}^{N} q_j \right) \mu_k + \sum_{j=k+1}^{N} q_j \mu_j + \sum_{j=1}^{N} \lambda_j \right. \\
&\quad \left. + \left(\sum_{j=k}^{N} q_j - s \right) \lambda_{T,k} + \sum_{j=1}^{k-1} q_j \lambda_{T,j} \right) \\
&= f(\mathbf{q}) \left(\left(s - \sum_{j=k+1}^{N} q_j \right) \left(\frac{1}{a_k} - 1 \right) \mu_k + \sum_{j=k+1}^{N} q_j \mu_j \left(\frac{1}{a_j} - 1 \right) \right. \\
&\quad \left. + \sum_{j=1}^{N} \lambda_j (a_j - 1) \right) - f(\mathbf{q}) \left(\left(\sum_{j=k}^{N} q_j - s \right) \left(1 - \frac{a_{k+1}}{a_k} \right) \lambda_{T,k} \right. \\
&\quad \left. + \sum_{j=1}^{k-1} q_j \lambda_{T,j} \left(1 - \frac{a_{j+1}}{a_j} \right) \right) \\
&\leq f(\mathbf{q}) \left(\sum_{j=1}^{N} \lambda_j (a_j - 1) - \left(\sum_{j=1}^{N} q_j - s \right) \right. \\
&\quad \left. \times \min_{1\leq j\leq N-1} \left\{ \left(1 - \frac{a_{j+1}}{a_j} \right) \lambda_{T,j} \right\} \right).
\end{aligned} \tag{4}$$

Note that the last expression in Eq. (4) is independent of k. Thus, Eq. (4) holds for $q_N < s$ and $\sum_{j=1}^{N} q_j \geq s$.

If $\sum_{j=1}^{N} q_j < s$, we have

$$\begin{aligned}
&\sum_{\mathbf{y}\geq 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) \\
&= \sum_{j=1}^{N} q_j \mu_j f(\mathbf{q}-\mathbf{e}(j)) + \sum_{j=1}^{N} \lambda_j f(\mathbf{q}+\mathbf{e}(j)) \\
&\quad - \left(\sum_{j=k+1}^{N} q_j \mu_j + \sum_{j=1}^{N} \lambda_j \right) f(\mathbf{q}) \\
&= f(\mathbf{q}) \left(\sum_{j=1}^{N} q_j \mu_j \left(\frac{1}{a_j} - 1 \right) + \sum_{j=1}^{N} \lambda_j (a_j - 1) \right) \\
&\leq \left(\sum_{j=1}^{N} \lambda_j (a_j - 1) \right) a_1^s.
\end{aligned} \tag{5}$$

Finally, we apply Theorem 1.18 in [3] to show that the Markov chain is ergodic. According to Theorem 1.18 in [3], we need to show that

$$\sum_{\mathbf{y}\geq 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) + \eta f(\mathbf{q}) \leq K \tag{6}$$

holds for all states for some constant $K \geq 0$ and $\eta > 0$.

For that purpose, we choose η satisfying $0 < \eta < \varepsilon$, where ε is defined in Eq. (2). Let $Y = \min_{1\leq j\leq N-1} \left\{ \left(1 - a_{j+1}/a_j \right) \lambda_{T,j} \right\}$ and $W = \eta + \sum_{j=1}^{N} \lambda_j (a_j - 1)$. We choose q^* as $q^* = s + W/Y$. Define

$$K = \max\left\{ 0, \max_{\left\{ \mathbf{q}: q_N < s, \sum_{j=1}^{N} q_j \leq q^* \right\}} \left\{ \sum_{\mathbf{y}\geq 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) + \eta f(\mathbf{q}) \right\} \right\}.$$

It is easy to see that K is non-negative and finite, i.e., $0 \leq K < \infty$.

For the chosen η and K, by their definitions, Eq. (6) holds if $q_N < s$ and $\sum_{j=1}^{N} q_j \le q^*$. If $q_N < s$ and $\sum_{j=1}^{N} q_j > q^*$, we have

$$\sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) + \eta f(\mathbf{q})$$
$$\le f(\mathbf{q})\left(\sum_{j=1}^{N}\lambda_j(a_j-1) - \left(\sum_{j=1}^{N} q_j - s\right)Y + \eta\right) \tag{7a}$$
$$= f(\mathbf{q})\left(W - \left(\sum_{j=1}^{N} q_j - s\right)Y\right) \tag{7b}$$
$$= f(\mathbf{q})\left((q^*-s)Y - \left(\sum_{j=1}^{N} q_j - s\right)Y\right) \tag{7c}$$
$$= f(\mathbf{q})Y\left(q^* - \sum_{j=1}^{N} q_j\right) \le 0,$$

where inequality (7a) follows the last inequality of Eq. (4), equality (7b) follows the definition of W, and equality (7c) follows the definition of q^*, respectively.

For $q_N \ge s$, Eq. (3) leads to

$$\sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) + \eta f(\mathbf{q})$$
$$= f(\mathbf{q})\left(h(a_N) + \sum_{j=1}^{N}\lambda_j(a_j - a_N) + \eta\right)$$
$$\le -f(\mathbf{q})(\varepsilon - \eta) \le 0.$$

Consequently, we have shown that Eq. (6) holds for all states. Therefore, by Theorem 1.18 in [3], the Markov chain is ergodic.

Proof of (1.2). To prove (1.2), instead of Theorem 1.18 in [3], part (c) of Theorem 1 in [4] is utilized. Note that the Lyapunov function $f(\mathbf{q})$ and constants $\{a_1, \ldots, a_N\}$ will be redefined for this case as well as in the proofs of cases (1.3) and (1.4).

First note that, for the function $h(x)$ defined in Eq. (1), we have $h(1) = h(x^*) = 0$, where $x^* = s\mu_N\left(\sum_{j=1}^{N}\lambda_j\right)^{-1} < 1$, and $h^{(1)}(1) = \lambda_1 + \cdots + \lambda_N - s\mu_N > 0$. Then we have $h(x) < 0$ for $x^* < x < 1$. Choose constants $\{a_1, \ldots, a_N\}$ satisfying the following conditions:

(1) $0 < x^* < a_N < a_{N-1} < \cdots < a_1 < 1$;

(2) $h(a_N) + \sum_{j=1}^{N}\lambda_j(a_j - a_N) \le 0.$

Define

$$f(\mathbf{q}) = \begin{cases} -\prod_{j=1}^{N} a_j^{q_j}, & q_j \ge 0,\ 1 \le j \le N; \\ 0, & \text{otherwise.}\end{cases}$$

According to part (c) of Theorem 1 in [4], the Markov chain is non-ergodic if:

$$\begin{aligned} &(1)\quad \sup_{\mathbf{q}\ge 0}\sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}}(f(\mathbf{q}) - f(\mathbf{y}))^+ < \infty; \\ &(2)\quad \sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) \ge 0, \quad \text{if } \mathbf{q} \in B; \\ &(3)\quad \exists\, \mathbf{q} \in B,\ f(\mathbf{q}) > \sup_{\mathbf{y}\notin B} f(\mathbf{y}), \end{aligned} \tag{8}$$

where $(x)^+ = \max\{0, x\}$ and B is a subset of states.

Condition (1) in Eq. (8) holds since $0 < a_1, \ldots, a_N < 1$ and $-1 \le f(\mathbf{q}) - f(\mathbf{y}) \le 1$ for all possible $\mathbf{q}$ and $\mathbf{y}$. Also note that, since $0 < a_j < 1$, we have $\lambda_{T,j} q_j a_j^{q_j} \to 0$ if $q_j \to \infty$. To verify conditions (2) and (3) in Eq. (8), we consider two cases: $q_N \ge s$ and $q_N < s$. If $q_N \ge s$, similarly to Eq. (3), we obtain

$$\sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) \ge -\left(\prod_{j=1}^{N} a_j^{q_j}\right)\left(h(a_N) + \sum_{j=1}^{N}\lambda_j(a_j - a_N)\right) \ge 0.$$

Let $Z = \sum_{j=1}^{N}\mu_j\left(1/a_j - 1\right)$ and $q^* = sY^{-1}Z + s$, where Y is defined in the proof of (1.1). If $q_N < s$ and $\sum_{j=1}^{N} q_j > q^*$, there exists k such that $\sum_{j=k+1}^{N} q_j < s$, $\sum_{j=k}^{N} q_j \ge s$, for $1 \le k \le N-1$. By the last equality in Eq. (4) and definitions of q^*, Y, and Z, we have

$$\sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y})$$
$$= \left(\prod_{j=1}^{N} a_j^{q_j}\right)\left(\left(\sum_{j=k+1}^{N} q_j - s\right)\left(\frac{1}{a_k} - 1\right)\mu_k - \sum_{j=k+1}^{N} q_j\mu_j\left(\frac{1}{a_j} - 1\right) + \sum_{j=1}^{N}\lambda_j(1 - a_j)\right)$$
$$+ \left(\prod_{j=1}^{N} a_j^{q_j}\right)\left(\left(\sum_{j=k}^{N} q_j - s\right)\left(1 - \frac{a_{k+1}}{a_k}\right)\lambda_{T,k} + \sum_{j=1}^{k-1} q_j\lambda_{T,j}\left(1 - \frac{a_{j+1}}{a_j}\right)\right)$$
$$\ge \left(\prod_{j=1}^{N} a_j^{q_j}\right)\left(\left(\sum_{j=1}^{N} q_j - s\right)\min_{1\le j\le N-1}\left\{\left(1 - \frac{a_{j+1}}{a_j}\right)\lambda_{T,j}\right\} - s\sum_{j=1}^{N}\mu_j\left(\frac{1}{a_j} - 1\right)\right)$$
$$= \left(\prod_{j=1}^{N} a_j^{q_j}\right)\left(\left(\sum_{j=1}^{N} q_j - s\right)Y - sZ\right)$$
$$\ge \left(\prod_{j=1}^{N} a_j^{q_j}\right)\left((q^* - s)Y - sZ\right) = 0.$$

Thus, condition (2) in Eq. (8) holds if $q_N < s$ and $\sum_{j=1}^{N} q_j > q^*$. Define

$$B = \{\mathbf{q} : q_N \ge s, \mathbf{q} \ge 0\} \cup \left\{\mathbf{q} : q_N < s \ \text{ and } \sum_{j=1}^{N} q_j > q^*, \mathbf{q} \ge 0\right\}.$$

Then condition (2) in Eq. (8) holds for any state $\mathbf{q}$ in the set B.

It is easy to see $\sup_{\mathbf{y}\in B} f(\mathbf{y}) = 0$. Since there are only a finite number of states not in the set B, it is easy to obtain $\sup_{\mathbf{y}\notin B} f(\mathbf{y}) < 0$. Then condition (3) in Eq. (8) holds for B. Therefore, the Markov chain is non-ergodic if $\lambda_1 + \cdots + \lambda_N > s\mu_N$.

Proof of (1.3). To prove (1.3), Theorem 1 in [4] is utilized again. Define

$$f(\mathbf{q}) = \begin{cases} \sum_{j=1}^{N} q_j, & q_j \ge 0,\ 1 \le j \le N; \\ 0, & \text{otherwise.}\end{cases} \tag{9}$$

According to part (a) of Theorem 1 in [4], the Markov chain is non-ergodic if:

$$\begin{aligned} &(1)\quad \sup_{\mathbf{q}\ge 0}\sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}}(f(\mathbf{q}) - f(\mathbf{y}))^+ < \infty; \\ &(2)\quad \sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) \ge 0, \quad \mathbf{q} \ge 0. \end{aligned} \tag{10}$$

Condition (1) in Eq. (10) holds since (i) $-1 \le f(\mathbf{q}) - f(\mathbf{y}) \le 1$ if $Q_{\mathbf{q},\mathbf{y}} \neq 0$ and (ii) $f(\mathbf{q}) - f(\mathbf{q} + \mathbf{e}(j+1) - \mathbf{e}(j)) = 0$. To verify condition (2) in Eq. (10), we consider the following three cases.

If $q_N \ge s$, similarly to Eq. (3), we obtain

$$\begin{aligned}\sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) &= \sum_{j=1}^{N} \lambda_j \left(1 + \sum_{i=1}^{N} q_i\right) + s\mu_N \left(\sum_{i=1}^{N} q_i - 1\right) \\ &\quad + \sum_{j=1}^{N-1} q_j \lambda_{T,j} \sum_{i=1}^{N} q_i - \left(s\mu_N + \sum_{j=1}^{N} \lambda_j + \sum_{j=1}^{N-1} q_j \lambda_{T,j}\right) \sum_{i=1}^{N} q_i \\ &= \sum_{j=1}^{N} \lambda_j - s\mu_N = 0. \end{aligned} \tag{11}$$

If $q_N < s$ and $\sum_{j=1}^{N} q_j \ge s$, there exists k such that $\sum_{j=k+1}^{N} q_j < s$ and $\sum_{j=k}^{N} q_j \ge s$, for $1 \le k \le N-1$. Similarly to Eq. (4), we have

$$\begin{aligned}\sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) &= \sum_{j=1}^{N} \lambda_j \left(1 + \sum_{i=1}^{N} q_i\right) + \left(s - \sum_{j=k+1}^{N} q_j\right) \\ &\quad \times \mu_k \left(\sum_{i=1}^{N} q_i - 1\right) + \sum_{j=k+1}^{N} q_j \mu_j \left(\sum_{i=1}^{N} q_i - 1\right) \\ &\quad + \left(\sum_{j=k}^{N} q_j - s\right) \lambda_{T,k} \sum_{i=1}^{N} q_i + \sum_{j=1}^{k-1} q_j \lambda_{T,j} \sum_{i=1}^{N} q_i \\ &\quad - \left(\sum_{j=1}^{N} \lambda_j + \left(s - \sum_{j=k+1}^{N} q_j\right) \mu_k + \sum_{j=k+1}^{N} q_j \mu_j \right. \\ &\quad \left. + \left(\sum_{j=k}^{N} q_j - s\right) \lambda_{T,k} + \sum_{j=1}^{k-1} q_j \lambda_{T,j}\right) \left(\sum_{i=1}^{N} q_i\right) \\ &= \sum_{j=1}^{N} \lambda_j - \left(s - \sum_{j=k+1}^{N} q_j\right) \mu_k - \sum_{j=k+1}^{N} q_j \mu_j \\ &\ge \sum_{j=1}^{N} \lambda_j - \left(s - \sum_{j=k+1}^{N} q_j\right) \mu_N - \sum_{j=k+1}^{N} q_j \mu_N \\ &= \sum_{j=1}^{N} \lambda_j - s\mu_N = 0. \end{aligned} \tag{12}$$

Note that the condition $\max_{1\le j\le N-1}\{\mu_j\} \le \mu_N$ is used in Eq. (12).

If $\sum_{j=1}^{N} q_j < s$, similarly to Eq. (5), we have

$$\begin{aligned}\sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) &= \sum_{j=1}^{N} \lambda_j \left(1 + \sum_{i=1}^{N} q_i\right) + \sum_{j=1}^{N} q_j \mu_j \left(\sum_{i=1}^{N} q_i - 1\right) \\ &\quad - \left(\sum_{j=1}^{N} \lambda_j + \sum_{j=1}^{N} q_j \mu_j\right) \left(\sum_{i=1}^{N} q_i\right) \\ &= \sum_{j=1}^{N} \lambda_j - \sum_{j=1}^{N} q_j \mu_j \\ &\ge \sum_{j=1}^{N} \lambda_j - \sum_{j=1}^{N} q_j \mu_N \quad \ge \sum_{j=1}^{N} \lambda_j - s\mu_N = 0. \end{aligned}$$

Condition (2) in Eq. (10) holds. Therefore, the Markov chain is non-ergodic.

Proof of (1.4). According to part (ii) of Theorem 2.2 in [13], the Markov chain is recurrent if

$$\sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) \le 0 \tag{13}$$

holds for all but a finite number of states $\mathbf{q}$, for the non-negative and unbounded function $f(\mathbf{q})$ defined in Eq. (9). To verify the condition in Eq. (13), we consider two cases. If $q_N \ge s$, Eq. (11) implies Eq. (13). If $q_N < s$ and $\sum_{j=1}^{N} q_j \ge s$, similarly to Eq. (12), we have, for all $1 \le k \le N-1$ such that $\sum_{j=k+1}^{N} q_j < s$ and $\sum_{j=k}^{N} q_j \ge s$,

$$\begin{aligned}\sum_{\mathbf{y}\ge 0} Q_{\mathbf{q},\mathbf{y}} f(\mathbf{y}) &= \sum_{j=1}^{N} \lambda_j - \left(s - \sum_{j=k+1}^{N} q_j\right) \mu_k - \sum_{j=k+1}^{N} q_j \mu_j \\ &\le \sum_{j=1}^{N} \lambda_j - \left(s - \sum_{j=k+1}^{N} q_j\right) \mu_N - \sum_{j=k+1}^{N} q_j \mu_N \\ &= \sum_{j=1}^{N} \lambda_j - s\mu_N = 0. \end{aligned} \tag{14}$$

Note that the condition $\min_{1\le j\le N-1}\{\mu_j\} \ge \mu_N$ is used in Eq. (14). Thus, Eq. (13) holds for all but states in the set $\{\mathbf{q} : q_1 + \cdots + q_N < s, \mathbf{q} \ge 0\}$, which only has a finite number of states. Consequently, the Markov chain is recurrent.

Acknowledgements

The authors would like to thank an anonymous referee for valuable comments and suggestions. The research was partially supported by NSF of China under grant 70325004.

References

[1] I.J.B.F. Adan, J. Wessels, W.H.M. Zijm, Analysis of the asymmetric shortest queue problem with threshold jockeying, Stochastic Models 7 (1991) 615–628.
[2] X. Chao, M. Miyazawa, M. Pinedo, Queueing Networks, Customers, Signals and Production Form Solutions, Wiley, Chichester, 1999.
[3] M.F. Chen, On three classical problems for Markov chains with continuous time parameters, Journal of Applied Probability 28 (2) (1991) 305–320.
[4] B.D. Choi, B. Kim, Non-ergodicity criteria for denumerable continuous time Markov processes, Operations Research Letters 32 (2004) 574–580.
[5] K.L. Chung, Markov Chains with Stationary Transition Probabilities, 2nd ed., Springer-Verlag, Berlin, 1967.
[6] J.W. Cohen, The Single Server Queue, North-Holland, Amsterdam, 1982.
[7] G. Fayolle, V.A Malyshev, M.V. Menshikov, Topics in the Constructive Theory of Countable Markov Chains, Cambridge University Press, 1995.
[8] F.G. Foster, On stochastic matrices associated with certain queueing processes, Annals of Mathematical Statistics 24 (1953) 355–360.
[9] B. Hajek, Optimal control of two interacting service stations, IEEE Transactions on Automatic Control AC-29 (1984) 491–499.
[10] S.A. Lippman, Applying a new device in the optimization of exponential queuing systems, Operations Research 23 (1975) 687–710.
[11] S.P. Meyn, R. Tweedie, Markov Chains and Stochastic Stability, Springer Verlag, 1993.
[12] H. Takagi, Queueing Systems, Vol. 1: Vacation and Priority Systems, Elsevier, Amsterdam, 1991.
[13] R. Tweedie, Sufficient conditions for regularity, recurrence and ergodicity of Markov processes, Mathematical Proceedings of the Cambridge Philosophical Society 78 (1975) 125–136.
[14] W. Whitt, Deciding which queue to join: Some counterexamples, Operations Research 34 (1986) 55–62.
[15] S.H. Xu, H. Chen, On the asymptote of the optimal routing policy for two service stations, IEEE Transactions on Automatic Control 38 (1990) 187–189.
[16] S.H. Xu, Y.Q. Zhao, Dynamic routing and jockeying controls in a two-station queueing system, Advances in Applied Probability 28 (1996) 1201–1226.
[17] Y. Zhao, W.K. Grassmann, Queueing analysis of a jockeying model, Operations Research 43 (1995) 520–529.