Queueing Analysis of a Jockeying Model

Yiqiang Zhao [1]
Department of Mathematics and Statistics
University of Winnipeg
Winnipeg, Manitoba
R3B 2E9 Canada

Winfried K. Grassmann
Department of Computational Science
University of Saskatchewan
Saskatoon, Saskatchewan
S7N 0W0 Canada

September 2, 2004

[1]Subject classification: Queues; multichannel; jockeying.

Abstract

In this paper, we solve a type of shortest queue problem, which is related to multibeam satellite systems. We assume that the packet interarrival times are independently distributed according to an arbitrary distribution function, that the service times are Markovian with possibly different service rates, that each of the servers has its own buffer for packet waiting, and that jockeying among buffers is permitted. Packets always join the shortest buffer(s). Jockeying takes place as soon as the difference between the longest and shortest buffers exceeds a pre-set number (not necessary 1). In this case, the last packet in a longest buffer jockeys instantaneously to the shortest buffer(s). We prove that the equilibrium distribution of packets in the system is modified vector-geometric. Expressions of main performance measures, including the average number of packets in the system, the average packet waiting time in the system and the average number of jockeying, are given. Based on the above solutions, numerical results are computed. By comparing the results for jockeying and non-jockeying models, we show that a significant improvement of the system performance is achieved for the jockeying model.

The performance study of a great number of satellite systems basically depends on the analysis of the related queueing systems. The major interesting measures of such analysis includes the system throughput, the average packet delay on the satellite, and the buffer overflow probability for the case of finite buffer size. Multibeam satellite systems have been studied extensively (for example, see Chlamtac and Ganz 1986, and Chang 1983), and it has been shown that they provide a greater system flexibility and a better performance. In such a system, all earth stations are organized into disjoint zones; packets generated from earth zones arrive at the satellite by using different possible access techniques, one or several buffers are provided at the satellite for the waiting packets to be processed or transmitted; and finally, the packets are sent to their destinations by the multi-down-link beams.

When there is more than one buffer on board, introducing jockeying of the waiting packets among the buffers seems to be a promising way to improve the performance of the systems. For example, if we allow a packet waiting in the buffer with many waiting packets to move to some other buffer with fewer waiting packets in it, then the average packet waiting time is obviously reduced. But the analysis of such jockeying systems is more difficult because we cannot deal with them by analysing only one specified input-output pair. Instead, we must handle the system as a whole.

According to different assumptions made on the system, a variety of different jockeying models could arise. In this paper, we consider a very general type of jockeying model, in which the following assumptions are made. The arriving packets follow a general process; that is, the time between any two successive arriving packets is described by an arbitrary non-negative random variable, the interarrival time. All the interarrival times are identically and independently distributed. There are several buffers in the system, each of them with an infinite capacity. An arrival packet always joins the shortest waiting line if it cannot be processed immediately. For any waiting line, the waiting packets are processed according to first come first served (FIFO) discipline. The processing time for any packet is a Markovian, that is, an exponential random variable. When the difference of the waiting packet numbers between the longest waiting line and the shortest one exceeds a certain threshold value, the last waiting packet is allowed to move (jockey) to the shortest waiting line.

Since Haight (1958) proposed and solved the shorter queue model (the shortest queue model with only two servers), the jockeying problem has been studied extensively, in particular

by Disney and Mitchell (1971), Elsayed and Bastani (1985), Kao and Lin (1990), Zhao and Grassmann (1990), Zhao (1990), and Adan, Wessels and Zijm (1991). Except Zhao and Grassmann, all authors considered only models with Markovian inputs. Among them, Kao and Lin solved the problem of jockeying as soon as the difference between queues exceeds one. They expressed their solution in terms of the eigenvalue of the rate matrix. Using the results of Kao and Lin, Zhao and Grassmann (1990) developed an explicit solution to the problem. They showed how to obtain certain initial probabilities for the system, and expressed the joint distributions of the queue lengths in the vector-geometric form. Recently, Nelson and Philips (1989) studied the response time for shortest queue routing by using approximations. As stated in their paper, shortest queue routing is a natural way to balance the load of a system across several processors and has been used as a load balancing mechanism as well as a scheduling mechanism in an effort to minimize job response time.

In this paper, we consider the general input model with a very flexible jockeying rule, in which the last packet in the longest queue jockeys to the shortest queue with an arbitrary probability distribution as soon as the difference of the waiting packet numbers between the longest queue and the shortest queue exceeds r, $r \geq 1$. Some special cases of this model have been considered by Zhao in his Ph.D thesis (Zhao 1990). After giving the definition of the model in the next section, we first consider the imbedded Markov chain of the model. We then obtain an explicit solution of the model, which also, as expected, has a vector-geometric form. Other interesting system measures are given, based on the probability distribution of the packet's number in the system. Numerical results are presented and analyzed. It turns out that significant improvements of the system performance can be achieved by allowing jockeying among the queues.

1 The Model and the Imbedded Markov Chain

In this section, we give the definition of the r difference jockeying problem with a general input, and point out some important properties of the transition probabilities.

In order to give a formal definition of the shortest queue model with r difference jockeying, we make the following assumptions:

a) The packets arrive singly with interarrival times identically and independently distributed

according to an arbitrary distribution function $A(t)$ with $A(t) = 0$ if $t < 0$, and they are not allowed to renege or balk.

b) There are c ($c \geq 2$) servers (for example, transponders or down-link beams), numbered $1, 2, \ldots, c$, in the system and each of them has its own buffer. In each buffer, service is rendered according to FIFO (first come first served) discipline. The c servers have independent exponential service times. The service times are independent of arrivals.

c) An arriving packet joins one of the shortest buffers with a pre-determined probability distribution.

d) Jockeying among the buffers is permitted. The last packet in the longest buffer(s) instantaneously jockeys to the shortest buffer(s) with a pre-determined probability distribution as soon as the difference of the packet numbers between the shortest buffer(s) and the longest buffer(s) exceeds r, $r \geq 1$.

A queueing system satisfying **a)** – **c)** is called the shortest queue model and denoted by $GI/(M/1)^c$. A shortest queue model satisfying **d)** is called the shortest queue model with r difference jockeying. When $r = 1$, we simply call it the shortest queue model with jockeying. When both the probability distributions mentioned in **c)** and **d)** are uniform, we call the shortest queue model symmetric; otherwise non-symmetric.

Let $X_k(t)$ represent the number of waiting packets in buffer k, $k = 1, 2, \ldots, c$, at time t, $t \geq 0$, including the packet in service, then $\{ \vec{X}(t) = (X_1(t), X_2(t), \ldots, X_c(t)); \ t \geq 0 \}$ is a stochastic process. The state space of this process can be described as

$$S = \{ \vec{i} = (i_1, \ldots, i_c) \mid i_j \ \text{ non-negative integer for } \ j = 1, 2, \ldots, c \\ \text{and } \ |i_k - i_l| \leq r \ \text{ for } \ k, l = 1, 2, \ldots, c \}.$$

The main purpose of this paper is to determine the limiting probabilities

$$\pi_{\vec{i}} = \lim_{t \to \infty} P\{ \vec{X}(t) = \vec{i} \}, \quad \vec{i} \in S,$$

when they exist.

In general, the above process is neither Markovian nor semi-Markovian. In order to analyze this model, we introduce the imbedded Markov chain for the system as follows. We obtain

an explicit formula for the probability distribution of buffer lengths (including the packets in service) for the imbedded Markov chain first. The probability distribution of the buffer lengths at a random time can then be found from the connection between a semi-Markov process and its imbedded Markov chain.

If t_l is the time just prior to the arrival of the lth packet, then $\{ \vec{X}_l = (X_1(t_l), X_2(t_l), \ldots, X_c(t_l)) ; \ l = 1, 2, \ldots \}$ is Markovian. Let $1/\lambda$ be the mean time between two successive arrivals, and let μ be the sum of all the service rates; that is, $\mu = \sum_{k=1}^{c} \mu_k$. The imbedded Markov chain $\{ \vec{X}_l ; \ l = 1, 2, \ldots \}$ formed in this way is ergodic if, and only if, the traffic intensity λ/μ is less than one. In the paper, we always assume that this is the case.

In the stable condition, the limiting or equilibrium probabilities

$$p_{\vec{i}} = \lim_{l \to \infty} P\{\vec{X}_l = \vec{i}\} , \quad \vec{i} \in S ,$$

exist and they are the same as the stationary probabilities of the imbedded Markov chain.

For any two states $\vec{i}$, $\vec{j} \in S$, the transition probability $p_{\vec{i}\vec{j}}$ can be found by conditioning on the interarrival time U_l; that is,

$$p_{\vec{i}\vec{j}} = \int_0^\infty P\{ \vec{X}_{l+1} = \vec{j} \mid U_l = t , \vec{X}_l = \vec{i} \} \, dA(t) .$$

An explicit determination of $p_{\vec{i}\vec{j}}$ can be obtained by using conditional probability arguments. Since only elementary algebraic manipulations are involved for the derivation and the final explicit expression of $p_{\vec{i}\vec{j}}$ is cumbersome, we will not produce it here. The readers, who are interested in details of deriving the explicit expression of $p_{\vec{i}\vec{j}}$, may refer to Zhao (1991), in which the same technique was used for obtaining the expression of $p_{\vec{i}\vec{j}}$ for the case of $r = 1$. Instead, we give the proofs, by using the same conditional probability argument, of the following properties of the transition probabilities. These properties are essential for proving our main results in this paper.

For a state $\vec{i}$, define $\#\vec{i}$ to be the number of packets in the system; that is, if $\vec{i} = (i_1, i_2, \ldots, i_c)$ then $\#\vec{i} = \sum_{k=1}^{c} i_k$, and define two functions $l(\cdot)$ and $s(\cdot)$ of states to be the number of packets in the longest buffer and the shortest buffer respectively; that is, $l(\vec{i}) = \max(i_1, i_2, \ldots, i_c)$ and $s(\vec{i}) = \min(i_1, i_2, \ldots, i_c)$. Define $\vec{1} = (1, 1, \ldots, 1)$.

Proposition 1 *Let $\vec{i}$, $\vec{j} \in S$ be two states in the state space. If $\#\vec{j} > \#\vec{i} + 1$, then $p_{\vec{i}\vec{j}} = 0$. If $\#\vec{j} = \#\vec{i} + 1$ and $s(\vec{j}) > 0$, then $p_{\vec{i}\vec{j}} = p_{\vec{i}+\vec{1}\ \vec{j}+\vec{1}}$, and*

$$\sum_{\substack{\vec{j}: \\ \#\vec{j}=\#\vec{i}+1}} p_{\vec{i}\vec{j}} = \beta_0 \,, \tag{1}$$

where

$$\beta_0 = \int_0^\infty e^{-\mu t} dA(t) \,. \tag{2}$$

If $\#\vec{j} < \#\vec{i} + 1$ with $l(\vec{j}) > r$, define k to be $(\#\vec{i} + 1) - \#\vec{j}$. Then $p_{\vec{i}\vec{j}} = p_{\vec{i}+\vec{1}\ \vec{j}+\vec{1}}$, and

$$\sum_{\substack{\vec{j}: \\ \#\vec{j}=\#\vec{i}+1-k}} p_{\vec{i}\vec{j}} = \beta_k \,, \tag{3}$$

where

$$\beta_k = \int_0^\infty \frac{(\mu t)^k}{k!} e^{-\mu t}\, dA(t) \,. \tag{4}$$

■

First, it is obvious that if $\#\vec{j} > \#\vec{i} + 1$, then $p_{\vec{i}\vec{j}} = 0$.

In the following we define $\vec{X}'_l$ to be the state immediately after the arrival of the lth packet. If $\#\vec{j} = \#\vec{i} + 1$ and $s(\vec{j}) > 0$, then

$$\begin{aligned}
p_{\vec{i}\vec{j}} &= \int_0^\infty P\{\ \vec{X}'_l = \vec{j}\ \mid\ \vec{X}_l = \vec{i}\ \} P\{\ \vec{X}_{l+1} = \vec{j}\ \mid\ U_l = t\,,\ \vec{X}'_l = \vec{j}\ \}\, dA(t) \\
&= \int_0^\infty P\{\ \vec{X}'_l = \vec{j}\ \mid\ \vec{X}_l = \vec{i}\ \} P\{\ \text{no packets served}\ \mid\ U_l = t\,,\ \text{all servers busy at}\ t_l\ \} dA(t) \\
&= \int_0^\infty P\{\ \vec{X}'_l = \vec{j} + \vec{1}\ \mid\ \vec{X}_l = \vec{i} + \vec{1}\ \} e^{-\mu t}\, dA(t) \\
&= p_{\vec{i}+\vec{1}, \vec{j}+\vec{1}} \,,
\end{aligned} \tag{5}$$

where $\vec{1} = (1, 1, \ldots, 1)$. Notice that whether $p_{\vec{i}\vec{j}} > 0$ or $p_{\vec{i}\vec{j}} = 0$ depends on whether $P\{\ \vec{X}'_l = \vec{j}\ \mid\ \vec{X}_l = \vec{i}\ \}$ is greater 0 or equal to 0. Also notice that the condition $s(\vec{j}) > 0$ is weaker than $s(\vec{i}) > 0$. It follows from (5) and $\sum_{\vec{j}} P\{\ \vec{X}'_l = \vec{j}\ \mid\ \vec{X}_l = \vec{i}\ \} = 1$ that

$$\sum_{\substack{\vec{j}: \\ \#\vec{j}=\#\vec{i}+1}} p_{\vec{i}\vec{j}} = \int_0^\infty e^{-\mu t}\, dA(t) = \beta_0 \,. \tag{6}$$

If $\#\vec{j} < \#\vec{i} + 1$ with $l(\vec{j}) > r$, define $k = (\#\vec{i} + 1) - \#\vec{j}$. If $\rightarrow \vec{i}_1 \rightarrow \vec{i}_2 \cdots \rightarrow \vec{i}_k$ is defined as the event that the system is in state $\vec{i}_1$ after the first packet served, in state $\vec{i}_2$ after the second packet served, ... , and in state $\vec{i}_k$ after the kth packet served, then, by using $l(\vec{j}) > r$, $p_{\vec{i}\vec{j}}$ becomes

$$\begin{aligned} p_{\vec{i}\vec{j}} &= \int_0^\infty \sum_{\vec{i}_1,\ldots,\vec{i}_k:E} P\{ \rightarrow \vec{i}_1 \cdots \rightarrow \vec{i}_k = \vec{X}_{l+1} = \vec{j} \mid U_l = t\,,\ \vec{X}_l = \vec{i}\,\}\, dA(t) \\ &= \int_0^\infty \sum_{\vec{i}_1,\ldots,\vec{i}_k:E} P\{ \rightarrow \vec{i}_1 + \vec{1} \cdots \rightarrow \vec{i}_k + \vec{1} = \vec{X}_{l+1} = \vec{j} + \vec{1} \mid U_l = t\,,\ \vec{X}_l = \vec{i} + \vec{1}\,\}\, dA(t) \\ &= p_{\vec{i}+\vec{1},\vec{j}+\vec{1}}\,, \end{aligned} \tag{7}$$

where E denotes the event $\{\#\vec{i}_1 = \#\vec{i},\ \#\vec{i}_2 = \#\vec{i} - 1,\ \ldots,\ \#\vec{i}_k = \#\vec{i} - k = \#\vec{j}\}$, and $\vec{1} = (1, 1, \ldots, 1)$.

It follows from (7) that

$$\begin{aligned} \sum_{\substack{\vec{j}: \\ \#\vec{j}=(\#\vec{i}+1)-k}} p_{\vec{i}\vec{j}} &= \int_0^\infty \sum_{\vec{i}_1,\ldots,\vec{i}_k:E} P\{ \rightarrow \vec{i}_1 \rightarrow \vec{i}_2 \cdots \rightarrow \vec{i}_k \ \mid\ U_l = t\,,\ \vec{X}_l = \vec{i}\,\} dA(t) \\ &= \int_0^\infty P\{\ k \text{ customers served } \mid\ U_l = t\,,\ \text{all servers busy at } \ t_l\ \} dA(t) \\ &= \int_0^\infty \frac{(\,\mu t\,)^k}{k!} e^{-\mu t}\, dA(t) \\ &= \beta_k\,. \end{aligned} \tag{8}$$

Notice that $l(\vec{j}) > 0$ is only a sufficient condition for the above property. One may prove (8) under a weaker condition.

2 Solution of the Imbedded Markov Chain

This section shows that the equilibrium probabilities of packet lengths for the imbedded Markov chain of the model under consideration obey a distribution which we call *vector-geometric*. By this, we mean that there is a constant, say ω, such that every state, except for boundary states, is related to exactly one other state in the sense that the ratio between their probabilities is ω. To do the proof, we introduce the concept of blocks of states, and the concept of groups of states. The proof exploits the special structure of the transition matrix.

Specifically, some states cannot be reached by an arrival. Moreover, after an arrival, the states in block $k+1$ can be reached from block k only through state $(k, k, \ldots, k)$. The proof also exploits the connection between the imbedded Markov chain of the $GI/(M/1)^c$ model with r difference jockeying and the imbedded Markov chain of the $GI/M/1$ model.

In order to state our main result, we need to partition the state space S into blocks according to the maximal number of packets in the buffers. Let

$$B_{<r} = \{\vec{i} \in S \ \mid \ l(\vec{i}) < r\}\,, \tag{9}$$

where r is the maximal allowable difference between the longest buffer and the shortest buffer, and

$$B_m = \{\vec{i} \in S \ \mid \ l(\vec{i}) = m\}\,, \quad m = r, r+1, \ldots\,. \tag{10}$$

The state space can be written as the union of the blocks:

$$S = B_{<r} \cup (\cup_{m=r}^{\infty} B_m)\,. \tag{11}$$

$B_{<r}$ and B_m are called blocks of states. It can be proved that the number of the states in $B_{<r}$ is r^c and that in B_m for all $m = r, r+1, \ldots$ is $(r+1)^c - r^c$. It follows from Proposition 1 that the transition matrix P of the imbedded Markov chain $\{\, \vec{X}_l \,,\ l = 1, 2, \ldots \}$ can be partitioned by blocks as follows.

$$P = \begin{bmatrix} A_{0\,0} & A_{0\,1} & & & & & & \\ A_{1\,0} & A_{1\,1} & A_0 & & & & & \\ A_{2\,0} & A_{2\,1} & A_1 & A_0 & & & & \\ A_{3\,0} & A_{3\,1} & A_2 & A_1 & A_0 & & & \\ A_{4\,0} & A_{4\,1} & A_3 & A_2 & A_1 & A_0 & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \ddots \end{bmatrix}. \tag{12}$$

Here, $A_{0\,0}$ is a submatrix of size $r^c \times r^c$, $A_{0\,1}$ of size $r^c \times [(r+1)^c - r^c]$, $A_{m\,0}$, $m = 1, 2, \ldots$, of size $[(r+1)^c - r^c] \times r^c$, and $A_{m\,1}$ for all $m = 1, 2, \ldots$ and A_m for all $m = 0, 1, 2, \ldots$ are of size $[(r+1)^c - r^c] \times [(r+1)^c - r^c]$.

The stationary equations can be written, according to blocks, as system I.

$$\left.\begin{array}{rcll} \vec{p}_{<r} & = & \vec{p}_{<r}A_{0\,0} + \sum_{k=1}^{\infty} \vec{p}_{k+r-1}A_{k\,0} & (a) \\ \vec{p}_{r} & = & \vec{p}_{<r}A_{0\,1} + \sum_{k=1}^{\infty} \vec{p}_{k+r-1}A_{k\,1} & (b) \\ \vec{p}_{m} & = & \sum_{k=0}^{\infty} \vec{p}_{k+m-1}A_{k}\,, \qquad m = r+1, r+2, \ldots . & (c) \end{array}\right\} \text{I}$$

The equations given in I(a) and I(b) are called boundary equations, and those in I(c) queue equations.

Main result: *The solution of the stationary probabilities of the imbedded Markov chain has a modified vector-geometric form. Specifically,*

$$\vec{p}_{m+1} = \sigma^c \vec{p}_m\,, \quad m = r+1, r+2, \ldots\,, \tag{13}$$

and $(\vec{p}_{<r}, \vec{p}_r, \vec{p}_{r+1})$ *can be obtained by solving the following equations*

$$(\vec{p}_{<r}, \vec{p}_r, \vec{p}_{r+1}) \begin{bmatrix} A_{0\,0} & A_{0\,1} & 0 \\ A_{1\,0} & A_{1\,1} & A_0 \\ \sum_{k=2}^{\infty} \sigma^{(k-2)c} A_{k\,0} & \sum_{k=2}^{\infty} \sigma^{(k-2)c} A_{k\,1} & \sum_{k=1}^{\infty} \sigma^{(k-1)c} A_k \end{bmatrix} = (\vec{p}_{<r}, \vec{p}_r, \vec{p}_{r+1})\,. \tag{14}$$

Here σ *is the unique solution for* x*, inside the unit circle, of the equation*

$$x = \sum_{k \geq 0} x^k \beta_k\,. \tag{15}$$

■

The proof of the main result consists of three parts. In part one, we prove that the queue equations have a vector-geometric solution. In part two, we prove that this vector-geometric solution for the queue equations is also satisfied by the boundary equations. In part three, we show that such a geometric parameter ω is unique and equal to σ^c.

Part one of the proof: The queue equations in I(c) have a vector-geometric solution

$$\vec{p}_{m+1} = \omega \vec{p}_m\,, \quad m = r, r+1, r+2, \ldots . \tag{16}$$

This is true due to the following Lemma and Theorem.

Lemma 1 *The queue equations in I(c) have a vector-geometric solution in the form of (16) if, and only if,* $\det(\omega I - \sum_{k=0}^{\infty} \omega^k A_k) = 0$.

Replacing $\vec{p}_{k+m-1}$ for all $k = 1, 2, \ldots$ in the queue equations by $\omega^k \vec{p}_{m-1}$ leads to

$$\omega \vec{p}_{m-1} = \sum_{k=0}^{\infty} \omega^k \vec{p}_{m-1} A_k \,. \tag{17}$$

Since this is a non-trivial solution, the determinant in question must be zero. ■

The imbedded Markov chain of the $GI/(M/1)^c$ queueing model with r difference jockeying, whose transition matrix is partitioned according to blocks, is a Markov chain of the $GI/M/1$ type defined in Neuts (1981). It follows from Lemma 1.2.4 by Neuts (1981) and the ergodicity assumption of the model that the rate matrix R has at least one positive eigenvalue ω_0 satisfying $0 < \omega_0 < 1$.

Theorem 1 *Let ω be an eigenvalue of the rate matrix R, then ω is a zero of the determinant* $\det(\omega I - \sum_{k=0}^{\infty} \omega^k A_k)$.

Proof: It follows from $R - \sum_{k=0}^{\infty} R^k A_k = 0$ and $\det(\omega I - R) = 0$ that

$$\begin{aligned}
&\det\left(\omega I - \sum_{k=0}^{\infty} \omega^k A_k\right) = \det\left[\left(\omega I - \sum_{k=0}^{\infty} \omega^k A_k\right) - \left(R - \sum_{k=0}^{\infty} R^k A_k\right)\right] \\
&= \det\left[(\omega I - R) - \sum_{k=0}^{\infty} (\omega^k I - R^k) A_k\right] \\
&= \det\left[(\omega I - R) - \sum_{k=1}^{\infty} (\omega^k I - R^k) A_k\right] \\
&= \det\left[(\omega I - R) - \sum_{k=1}^{\infty} (\omega I - R)(\omega^{k-1} I + \omega^{k-2} R + \cdots + \omega R^{k-2} + R^{k-1}) A_k\right] \\
&= \det\left\{(\omega I - R)\left[I - \sum_{k=1}^{\infty} (\omega^{k-1} I + \omega^{k-2} R + \cdots + \omega R^{k-2} + R^{k-1}) A_k\right]\right\} \\
&= \det(\omega I - R) \det\left[I - \sum_{k=1}^{\infty} (\omega^{k-1} I + \omega^{k-2} R + \cdots + \omega R^{k-2} + R^{k-1}) A_k\right] \\
&= 0 \,.
\end{aligned}$$

■

The above Lemma and Theorem guarantee that the determinant $\det(\omega I - \sum_{k=0}^{\infty} \omega^k A_k)$ has at least one zero ω_0 satisfying $0 < \omega_0 < 1$. This means that there is at least one $\omega = \omega_0$

such that the queue equations have a vector-geometric solution in the form of that given in (16).

Part two of the proof: The vector-geometric solution for the queue equations is not necessarily satisfied by the boundary equations. In this part, we show that starting from the last element $p_{(r,r,\ldots,r)}$ in $\vec{p}_r$, the solution given in (16) can be satisfied by the boundary equations. Specifically, we prove that

$$p_{(r+1,r+1,\ldots,r+1)} = \omega p_{(r,r,\ldots,r)}\,, \tag{18}$$

$$\vec{p}_{m+1} = \omega \vec{p}_m\,, \qquad m = r+1, r+2, \ldots\,, \tag{19}$$

also satisfy the boundary equations. Replacing the above two expressions into the stationary equations given in system I gives us system II with $X_\# = r^c + [(r+1)^c - r^c] + [(r+1)^c - r^c]$ unknowns and $E_\# = X_\# + [(r+1)^c - r^c] + 1$ equations.

$$\left.\begin{array}{lcll}
\vec{p}_{<r} & = & \vec{p}_{<r}A_{0\,0} + \vec{p}_r A_{1\,0} + \vec{p}_{r+1}\sum_{k=2}^{\infty}\omega^{k-2}A_{k\,0} & (a)\\
\vec{p}_r & = & \vec{p}_{<r}A_{0\,1} + \vec{p}_r A_{1\,1} + \vec{p}_{r+1}\sum_{k=2}^{\infty}\omega^{k-2}A_{k\,1} & (b)\\
\vec{p}_{r+1} & = & \vec{p}_r A_0 + \vec{p}_{r+1}\sum_{k=1}^{\infty}\omega^{k-1}A_k & (c)\\
\omega\vec{p}_{r+1} & = & \vec{p}_{r+1}\sum_{k=0}^{\infty}\omega^{k}A_k & (d)\\
p_{(r+1,r+1,\ldots,r+1)} & = & \omega p_{(r,r,\ldots,r)}\,. & (e)
\end{array}\right\}\ \text{II}$$

Our aim is to prove that system II has a non-trivial solution for $\vec{p}_{<r}$, $\vec{p}_r$ and $\vec{p}_{r+1}$. If we can prove that there are $[(r+1)^c - r^c] + 1$ equations in system II which are redundant, then we are done.

First, we prove that all equations in II(d) are redundant. Since only the last row of A_0 is non-zero, $\vec{p}_{r+1}A_0$ can be written as $\omega\vec{p}_r A_0$ due to equation II(e). Therefore, II(d), which is II(c) multiplied by ω, is redundant. Next, we prove that II(e) is redundant. In order to do so, add all boundary equations in II(a) and II(b) together to form one equation, and notice the property of the stochastic matrix P; that is,

$$A_{0\,0}\vec{e} + A_{0\,1}\vec{e} = I\vec{e}, \tag{20}$$

$$(A_{k\,0} + A_{k\,1})\vec{e} = I\vec{e} - (A_0 + A_1 + \cdots + A_{k-1})\vec{e}, \quad k = 1, 2, \ldots, \tag{21}$$

where the transpose of $\vec{e}$ is $(1, 1, \ldots, 1)$ with a proper size. This leads to

$$\vec{p}_r A_0 \vec{e} = \vec{p}_{r+1} \sum_{k=2}^{\infty} \omega^{k-2} (I - A_0 - A_1 - \cdots - A_{k-1}) \vec{e}. \tag{22}$$

Add all equations corresponding to block B_{r+1} or all equations in II(c) together, which gives

$$\vec{p}_{r+1} \vec{e} = \vec{p}_r A_0 \vec{e} + \vec{p}_{r+1} \sum_{k=1}^{\infty} \omega^{k-1} A_k \vec{e}. \tag{23}$$

Now, combining (22) and (23) gives

$$\vec{p}_{r+1} A_0 \vec{e} = \vec{p}_{r+1} \sum_{k=2}^{\infty} \omega^{k-1} (I - A_0 - A_1 - \cdots - A_{k-1}) \vec{e}. \tag{24}$$

And (22) and (24) lead to

$$(\vec{p}_{r+1} - \omega \vec{p}_r) A_0 \vec{e} = 0\,. \tag{25}$$

Since only the last row of A_0 is non-zero, the sum of which equals β_0, (25) becomes

$$(p_{(r+1,\ldots,r+1)} - \omega p_{(r,\ldots,r)}) \beta_0 = 0\,. \tag{26}$$

This gives $p_{(r+1,\ldots,r+1)} = \omega p_{(r,\ldots,r)}$. Hence, equation II(e) is redundant.

Part three of the proof: It is clear, according to the proof in part two, that ω is unique. Otherwise, for each ω we have a solution for the system, which contradicts the uniqueness of the solution of the system. The only thing left now is to prove that $\omega = \sigma^c$ where σ is the unique solution, inside of the unit circle, of equation (15). For this purpose, we define groups G_k of states as follows:

$$G_k = \{\vec{i} \in S \mid \#\vec{i} = k\}\,, \quad k = 0, 1, 2, \ldots\,. \tag{27}$$

Therefore, group G_k consists of all states $\vec{i}$, which contain the same number k of packets. Add up all stationary equations corresponding to groups G_0, G_1, $\ldots$, G_{rc-1} to form one equation, and add up all equations corresponding to group G_{rc+k} to form one equation for each $k = 0, 1, 2, \ldots$. Denoted by p_k the sum of all stationary probabilities over group G_k: $\sum_{\#\vec{i}=k} p_{\vec{i}}$ and notice the property of the transition probabilities given in Proposition 1; that is, the sum of all transition probabilities over a group is the constant β_k, the resulting equations are given as

$$p_{rc} = \sum_{k=0}^{\infty} p_{rc+k} (1 - \beta_0 - \beta_1 - \cdots - \beta_k) \tag{28}$$

$$p_{rc+m} = \sum_{k=0}^{\infty} p_{rc+m+k-1} \beta_k\,, \quad m = 1, 2, \ldots\,. \tag{29}$$

This is a $GI/M/1$ queueing system. Therefore

$$p_{m+rc+1} = \sigma p_{m+rc}\,, \quad m = 0, 1, 2, \ldots\,. \tag{30}$$

On the other hand, it follows from

$$\vec{p}_{m+1} = \omega \vec{p}_m\,, \quad m = r+1, r+2, \ldots, \tag{31}$$

that

$$p_{m+(r+1)c} = \omega p_{m+rc} \tag{32}$$

at least for $m = 1, 2, \ldots$. In fact, it is valid also for $m = 0$ due to the equation II(e). For our purpose here, it is enough to know that there is an m_0 such that (32) is valid for all $m \geq m_0$. Combining (30) and (32) gives that

$$\sigma^c p_{m+rc} = \omega p_{m+rc} \tag{33}$$

for all $m = m_0, m_0+1, \ldots$. The only possibility is that $\omega = \sigma^c$. This completes the whole proof of our main result.

3 Solution of the Model

In the previous section, we have seen that the solution of the imbedded Markov chain has a modified vector-geometric solution. We now show that this is also the case for the solution of the equilibrium probabilities of the buffer lengths at a random time.

If $X_i(t)$, $i = 1, 2, \ldots, c$, represents the lengths of buffers at time $t \geq 0$, then $\{\, \vec{X}(t) = (\, X_1(t), X_2(t), \ldots, X_c(t)\,)\ ;\ t \geq 0\,\}$ describes a stochastic process. Generally speaking, this process is neither Markovian nor semi-Markovian. In this section, we show that the equilibrium probabilities of the buffer lengths at a random time

$$\pi_{\vec{i}} = \lim_{t\to\infty} P\{\, \vec{X}(t) = \vec{i}\,\}\,, \quad \vec{i} \in S\,,$$

also have a modified vector-geometric solution with the same parameter $\omega = \sigma^c$. For a proof, we use the fact that the equilibrium probabilities of the imbedded Markov chain are the same as that of the imbedded semi-Markov chain. Define $X_i^*(t)$, $i = 1, 2, \ldots, c$, to be the lengths of buffers just prior to the last arrival before time t, then $\{\, \vec{X}^*(t) =$

$(X_1^*(t), X_2^*(t), \ldots, X_c^*(t)), \ t \geq 0 \}$ is a semi-Markov chain which is called the imbedded semi-Markov chain of the $GI/(M/1)^c$ model with r difference jockeying.

The main theorem in the section that the equilibrium probabilities of the buffer lengths at a random time have a modified vector-geometric solution is proved by using the formula

$$\begin{aligned}\pi_{\vec{j}} &= \sum_{\vec{i}} p_{\vec{i}} \int_0^\infty P\{ \vec{X}(t_n + t) = \vec{j} \mid U_n > t, \ \vec{X}(t_n) = \vec{i} \}\lambda[1 - A(t)]dt \\ &= \sum_{\vec{i}} p_{\vec{i}} \int_0^\infty P\{ \vec{X}(t_{n+1}) = \vec{j} \mid U_n = t, \ \vec{X}(t_n) = \vec{i} \}\lambda[1 - A(t)]dt .\end{aligned}$$

For a proof of the first equality, use fact (iv) on page 351 of Gross and Harris and the fact that the limiting probability distribution of the imbedded semi-Markov chain is the same as that of the imbedded Markov chain since the mean time spent in every state during a visit is the same equal to $1/\lambda$. The proof of the second equality is based on the memorylessness of the service times and the independence of the service times and the interarrival times. Under these conditions, the transition probability from $\vec{i}$ at the previous imbedded epoch to $\vec{j}$ is independent of t the length of the time.

Partition the equilibrium probability vector $\vec{\pi}$ according to blocks defined on the state space: $\vec{\pi} = (\vec{\pi}_{<r}, \vec{\pi}_r, \vec{\pi}_{r+1}, \ldots\ldots)$.

Theorem 2 *For the shortest queue model $GI/(M/1)^c$ with r difference jockeying, the equilibrium probability vectors $\vec{\pi}_m$, $m \geq r+1$, have a solution in the modified vector-geometric form,*

$$\vec{\pi}_{m+1} = \vec{\pi}_m \sigma^c , \quad m = r+1, r+2, \ldots ,$$

where σ is the unique solution for x, inside the unit circle, of equation (15).

Proof: Since

$$\begin{aligned} & P\{ \vec{X}(t_{n+1}) = \vec{j} + \vec{1} \mid U_n = t, \ \vec{X}(t_n) = \vec{i} + \vec{1} \} \\ = \ & P\{ \vec{X}(t_{n+1}) = \vec{j} \mid U_n = t, \ \vec{X}(t_n) = \vec{i} \}\end{aligned}$$

and

$$p_{\vec{i}+\vec{1}} = p_{\vec{i}} \sigma^c$$

for $\vec{i} \in B_m$ and $\vec{j} \in B_m$ with $m \geq r+1$,

$$\begin{aligned}
\pi_{\vec{j}+\vec{1}} &= \sum_{\vec{i}+\vec{1}} p_{\vec{i}+\vec{1}} \int_0^\infty P\{\, \vec{X}(t_{n+1}) = \vec{j}+\vec{1} \;\mid\; U_n = t\,,\; \vec{X}(t_n) = \vec{i}+\vec{1}\,\}\, \lambda[\,1-A(t)\,]dt \\
&= \sum_{\vec{i}+\vec{1}} \sigma^c p_{\vec{i}} \int_0^\infty P\{\, \vec{X}(t_{n+1}) = \vec{j} \;\mid\; U_n = t\,,\; \vec{X}(t_n) = \vec{i}\,\}\lambda[\,1-A(t)\,]dt \\
&= \sigma^c \sum_{\vec{i}} p_{\vec{i}} \int_0^\infty P\{\, \vec{X}(t_{n+1}) = \vec{j} \;\mid\; U_n = t\,,\; \vec{X}(t_n) = \vec{i}\,\}\lambda[\,1-A(t)\,]dt \\
&= \sigma^c \pi_{\vec{j}}\,.
\end{aligned}$$

■

4 The Model with Markovian Input

If the arrivals are Markovian, the stochastic process $\vec{X}(t)$ defined in Section 1 is also Markovian. Instead of considering the imbedded Markov chain, we can directly work with the continuous time Markov chain $\vec{X}(t)$.

Let Q be the generator of the Markov chain $\vec{X}(t)$, and let the stationary probability vector $\vec{p}$ be partitioned according to blocks. Then

$$Q = \begin{bmatrix}
R_{0\,0} & R_{0\,1} & & & & \\
R_{1\,0} & R_{1\,1} & Q_0 & & & \\
 & Q_2 & Q_1 & Q_0 & & \\
 & & Q_2 & Q_1 & Q_0 & \\
 & & & \vdots & \vdots & \cdots\cdots \\
 & & & & &
\end{bmatrix}, \tag{34}$$

and the stationary equations can be written as

$$0 = \vec{p}_0 R_{0\,0} + \vec{p}_1 R_{1\,0}\,, \tag{35}$$

$$0 = \vec{p}_0 R_{0\,1} + \vec{p}_1 R_{1\,1} + \vec{p}_2 Q_2\,, \tag{36}$$

$$0 = \vec{p}_{j-1} Q_0 + \vec{p}_j Q_1 + \vec{p}_{j+1} Q_2\,, \quad j \geq 2\,. \tag{37}$$

If we repeat what we did in Section 3, then we have the following main result.

Main result: *The solution of the stationary probabilities for the $M/(M/1)^c$ queueing model with r difference jockeying has a modified vector-geometric form. Specifically,*

$$\vec{p}_{m+1} = \rho^c \vec{p}_m \,, \quad m \geq r+1 \tag{38}$$

and $(\vec{p}_{<r}, \vec{p}_r, \vec{p}_{r+1})$ is determined by solving the following equations

$$(\vec{p}_{<r}, \vec{p}_r, \vec{p}_{r+1}) \begin{bmatrix} R_{0\,0} & R_{0\,1} & 0 \\ R_{1\,0} & R_{1\,1} & Q_0 \\ 0 & Q_2 & Q_1 + \rho^c Q_2 \end{bmatrix} = 0 \,, \tag{39}$$

where $\rho = \lambda/\mu$ is the traffic intensity of the system.

Example: The symmetric $M/(M/1)^2$ model with $r = 2$ difference jockeying. In this case,

$$B_{<2} = \{(0,0),(1,0),(0,1),(1,1)\}, \tag{40}$$

$$B_2 = \{(2,0),(2,1),(0,2),(1,2),(2,2)\} \tag{41}$$

and

$$B_{m+1} = \{\vec{i} + (1,1) \mid \vec{i} \in B_m\} \,, \quad m = 2,3,4,\ldots \,. \tag{42}$$

The generator of the Markov chain is partitioned according to blocks as that in (34) with

$$R_{0\,0} = \begin{bmatrix} -\lambda & \lambda/2 & \lambda/2 & 0 \\ \mu_1 & -(\lambda+\mu_1) & 0 & \lambda \\ \mu_2 & 0 & -(\lambda+\mu_2) & \lambda \\ 0 & \mu_2 & \mu_1 & -(\lambda+\mu) \end{bmatrix} , \tag{43}$$

$$R_{0\,1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda/2 & 0 & \lambda/2 & 0 \end{bmatrix} , \tag{44}$$

$$R_{1\,0} = \begin{bmatrix} 0 & \mu_1 & 0 & 0 \\ 0 & 0 & 0 & \mu_1 \\ 0 & 0 & \mu_2 & 0 \\ 0 & 0 & 0 & \mu_2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{45}$$

$$R_{1\,1} = \begin{bmatrix} -(\lambda+\mu_1) & \lambda & 0 & 0 & 0 \\ \mu_2 & -(\lambda+\mu) & 0 & 0 & \lambda \\ 0 & 0 & -(\lambda+\mu_2) & \lambda & 0 \\ 0 & 0 & \mu_1 & -(\lambda+\mu) & \lambda \\ 0 & \mu_2 & 0 & \mu_1 & -(\lambda+\mu) \end{bmatrix}, \tag{46}$$

$$Q_0 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda/2 & 0 & \lambda/2 & 0 \end{bmatrix}, \tag{47}$$

$$Q_1 = \begin{bmatrix} -(\lambda+\mu) & \lambda & 0 & 0 & 0 \\ \mu_2 & -(\lambda+\mu) & 0 & 0 & \lambda \\ 0 & 0 & -(\lambda+\mu) & \lambda & 0 \\ 0 & 0 & \mu_1 & -(\lambda+\mu) & \lambda \\ 0 & \mu_2 & 0 & \mu_1 & -(\lambda+\mu) \end{bmatrix}, \tag{48}$$

and

$$Q_2 = \begin{bmatrix} 0 & \mu & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu_1 \\ 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & \mu_2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} . \tag{49}$$

According to the main result,

$$\vec{p}_{m+1} = \rho^c \vec{p}_m \,, \quad m \geq 3 \,, \tag{50}$$

and $(\vec{p}_{<2}, \vec{p}_2, \vec{p}_3)$ is obtained by solving the equations given in (39). A numerical example is given in Figure 2.

For the example here, we can further express all stationary probabilities $p_{\vec{i}}$ for $\vec{i} \in \cup_{m=1}^{\infty} B_{2+m}$ in terms of only the probability $p_{(2,2,\ldots,2)}$. To show that, let us examine the first one and the third one of the queue equations in any block. They are corresponding to the states, which cannot be reached by an arrival. These two equations give the relationships between $p_{3\,1}$ and $p_{3\,2}$, and $p_{1\,3}$ and $p_{2\,3}$ respectively. This makes it easy to solve the first four equations explicitly and to express all $p_{3\,1}$, $p_{3\,2}$, $p_{1\,3}$ and $p_{2\,3}$ in terms of $p_{2\,2}$. The probabilities $p_{\vec{i}}$ for $\vec{i} \in \cup_{m \geq 3} B_m$ can then be expressed in terms of only $p_{(2,2,\ldots,2)}$:

$$p_{k+3\,k+3} = \rho^{2(k+1)} p_{2\,2} \,, \tag{51}$$

$$p_{k+1\,k+3} = \rho^{2k} C_1 p_{2\,2} \,, \tag{52}$$

$$p_{k+2\,k+3} = \rho^{2k} \left(\frac{\lambda + \mu}{\mu_2} \right) C_1 p_{2\,2} \,, \tag{53}$$

$$p_{k+3\,k+1} = \rho^{2k} C_2 p_{2\,2} \,, \tag{54}$$

$$p_{k+3\,k+2} = \rho^{2k} \left(\frac{\lambda + \mu}{\mu_1} \right) C_2 p_{2\,2} \,, \tag{55}$$

$k = 0, 1, 2, \ldots$, where

$$C_1 = \frac{(\lambda/2) + \mu_1 \rho^2}{\left[\frac{(\lambda+\mu)^2}{\mu_1} - (\lambda + \mu\rho^2) \right]} \tag{56}$$

and

$$C_2 = \frac{(\lambda/2) + \mu_2 \rho^2}{\left[\frac{(\lambda+\mu)^2}{\mu_2} - (\lambda + \mu\rho^2) \right]} \,. \tag{57}$$

Our results agree with the ones obtained by Adan *et al.* (1991). Our proof could be potentially used to generalize the above expressions to the $GI/(M/1)^c$ model with r difference jockeying.

5 Numerical Results and Performance Analysis

The numerical results given in this section show that the introduction of jockeying can significantly improve the system performance. The performance measures used are L, the average number of packets in the system, and W, the average time a packet spends in the system. The results of section 3 allow us to find these measures as follows.

Consequence 1 *Let* $p_k = \sum_{\vec{i}:\, \#\vec{i}=k} p_{\vec{i}}$ *for* $k \geq 1$. *Then*

$$\begin{aligned} L &= \sum_{k=1}^{rc-1} k p_k + \frac{1}{(1-\sigma^c)} \sum_{k=rc}^{(r+1)c-1} \left(k + \frac{c\sigma^c}{1-\sigma^c}\right) p_k\,, \quad \text{and} \\ W &= \frac{L}{\lambda}. \end{aligned}$$

■

Therefore, the average number L of packets in the system and the average waiting time W of a packet in the system depend on σ and the probability vectors $\vec{p}_{<r}$, $\vec{p}_r$ and $\vec{p}_{r+1}$. Since σ , the root of the characteristic equation of the $GI/M/c$ model, is easy to find (see for example Chaudhry, Agarwal and Templeton (1992)), the only problem left is the determination of the probability vectors $\vec{p}_{<r}$, $\vec{p}_r$ and $\vec{p}_{r+1}$. One may use different numerical methods to compute these vectors. Here, we use the GTH-algorithm given in Grassmann Taksar and Heyman (1985) to obtain numerical results.

As we mentioned in the introduction, jockeying shortens the average packet waiting time in the system. This improvement is significant when service rates are uneven. Since the average number of packets in the system is proportional to the average packet waiting time in the system, jockeying also decreases this number. For brevity, we restrict the discussion to Markovian input models with two servers.

For all values of the traffic intensity ρ, the average packet waiting time will be shortened when jockeying is allowed. A significant improvement of the average packet waiting time would be expected if service times are uneven (see Table 1 to Table 3). But, under very

light or heavy traffic conditions, the improvement of the average packet waiting time would be offsetted by the cost of jockeying. Therefore, the system performance may not be actually improved. Generally speaking, for almost all interesting traffic intensities ρ, say 0.25 to 0.95, the performance of the system would be improved when jockeying is allowed (see Table 4 and Table 5). If the ratio between service rates is 9 to 1, the average waiting time in the system is shortened by more than 20 %. Table 1 gives the average waiting time W for $\rho = 0.7$, for $r = 1$, 2, and the non-jockeying model, and for a ratio between service times which is 9 to 1, The percentages of the improvement of the average waiting time for jockeying models compared to the non-jockeying model are also given.

Table 1: The average packet waiting time W and the improvement for jockeying models.

	W		
	9:1	3:1	1:1
r=1	2.210	2.051	1.961
improvement (%)	38.3	17.6	7.0
$r = 2$	2.667	2.273	2.062
improvement (%)	25.6	8.7	2.2
non-jockeying	3.582	2.489	2.108

In Table 2 and Table 3, the percentages of the improvement of the average packet waiting time W and the average number of packets in the system L for jockeying models compared to non-jockeying models are given for different values of the traffic intensity ρ, different maximal values r between buffer lengths, and different ratios of service rates.

Table 2: The performance improvement when the ratio of the service rates is 9 to 1.

ρ	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
$r = 1$ (%)	22.9	29.4	34.1	37.1	38.3	37.0	30.5	21.7
$r = 2$ (%)	4.67	9.77	15.58	21.12	25.57	27.88	25.25	18.71
$r = 3$ (%)	0.63	2.38	5.78	10.55	15.81	20.15	20.53	15.94

Table 3: The performance improvement when the ratio of the service rates is 3 to 1.

ρ	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
$r = 1$ (%)	9.1	12.6	15.3	17.0	17.6	16.3	11.9	7.4
$r = 2$ (%)	1.13	2.61	4.60	6.78	8.67	9.53	7.95	5.29
$r = 3$ (%)	0.11	0.42	1.14	2.32	3.83	5.17	5.09	3.64

The improvement of the system performance due to jockeying is particularly pronounced in models with uneven service rates. In our numerical experiments, different situations with ratios of service rates μ_1 to μ_2 being 9 and 3 to 1 respectively, are compared. The average packet waiting time is decreased by more than 38% for highly uneven service rates with service ratio 9 to 1 and a moderate traffic intensity $\rho = 0.7$. In contrast to the cases with uneven service rates, the average packet waiting time in the system with equal service rates can only be decreased by a maximum 7.2% (see Figure 4). Since the average number of the packets in the system is linearly proportional to the average packet waiting time in the system, this number is be decreased by the same percentages.

Another interesting performance measure is the average times of jockeying of a packet or the average times that a packet changes buffers, which is denoted by J_N. This number decreases as r increases. It is not difficult to derive a formula for J_N by using the main result and conditioning on the system states. Notice that only for those states $\vec{i} = (i_1, i_2)$ with $\min(i_1, i_2) \geq 1$ and $|i_1 - i_2| = r$, a service completion from the shorter queue contributes to J_N. Specifically,

$$J_N = \frac{1}{\lambda} \sum_{k=1}^{\infty} (\mu_1 p_{k,\,k+r} + \mu_2 p_{k+r,\,k}) \tag{58}$$

$$= \left(\frac{1}{\lambda}\right) \left(\frac{1}{1-\rho^c}\right) (\mu_1 p_{1,\,1+r} + \mu_2 p_{1+r,\,1}) \,. \tag{59}$$

In Table 4 and 5, where $\rho = 0.7$, the average times of jockeying of a packet and the improvement of the average waiting time of a packet for jockeying models compared to non-jockeying models are given and compared for different values of r with the ratio of service

rates 9 to 1 and 3 to 1 respectively.

Table 4: J_N and the improvement of W for a ratio of 9 to 1 between service rates.

r	1	2	3	4	5	∞
J_N	.2414	.0860	.0369	.0163	.0073	0
$ImpW(\%)$	38.3	25.6	15.8	9.2	5.0	0

Table 5: J_N and the improvement of W for a ratio of 3 to 1 between service rates.

r	1	2	3	4	5	∞
J_N	.1957	.0459	.0136	.0044	.0015	0
$ImpW$ (%)	17.6	8.7	3.8	1.6	0.6	0

In the above tables, we can see that the improvement of the average waiting time of a packet for a jockeying model compared to a non-jockeying model is as large as 38.3 percent and the mean number of a packet jockeys is below 0.25.

6 Conclusions

In this paper, we analyzed the $GI/(M/1)^c$ jockeying model and obtained a modified vector-geometric solution for the equilibrium probabilities of the number of packets in the system. The expressions of the average number of packets in the system, the average packet waiting time in the system and the average times of jockeying of a packet are also given. Numerical solutions of the model only depend on the root of the characteristic equation of the model and the probability vectors $\vec{p}_{<r}$, $\vec{p}_r$ and $\vec{p}_{r+1}$. The root can be found by using root finding method and the probability vectors $\vec{p}_{<r}$, $\vec{p}_r$ and $\vec{p}_{r+1}$ can be numerically obtained by using different methods. Significant improvements of the average number of packets in the system, and the average packet waiting time in the system are shown for the jockeying models with uneven service rates.

Acknowledgement

Y. Zhao acknowledges that this research was supported by a grant from the Canadian Institute for Telecommunications Research under the NCE program of the Government of Canada when he was with Queen's University and W.K. Grassmann acknowledges the research support through a research grant of the NSERC. The both authors gratefully thank Professor L.L. Campbell for reading the manuscript and suggesting a comparison between jockeying and non-jockeying models, and referees for their valuable suggestions and comments which improved the presentation of the paper.

References

Adan, I.J.B.F., J. Wessels and W.H.M. Zijm. 1991. Analysis of the asymmetric shortest queue problem with threshold jockeying. *Stochastic Models* **7**(41), 615–28.

Chang, J.F. 1983. A multibeam packet satellite using random access techniques. *IEEE Trans. Commun.* **COM-31**, 1143–1154.

Chaudhry, M.L., M. Agarwal and J.G.C. Templeton. 1992. Exact and approximate numerical solutions of steady-state distributions arising in the queue $GI/G/1$. *Queueing Systems* **10**, 105–152.

Chlamtac, I. and O. Ganz. 1986. Performance of multibeam packet satellite systems with conflict free scheduling. *IEEE Trans. Commun.* **COM-34**, 1016–1023.

Disney, R.L. and W.E. Mitchell. 1971. A solution for queues with instantaneous jockeying and other customer selection rules. *Naval Research Logistics* **17**, 315–325.

Elsayed, E.A. and A. Bastani. 1985. General solutions of jockeying problem. *European Journal of Operational Research* **22**, 387–396.

Gross, D. and C. Harris. 1985. *Fundamentals of Queueing Theory.* John Wiley And Sons, New York.

Haight, F.A. 1958. Two queues in parallel. *Biometrika* **45**, 401–410.

Kao, E.P.C. and C. Lin. 1990. A matrix-geometric solution of the jockeying problem. *European Journal of Operational Research* **44**, 67–74.

Nelson, R.D. and T.K. Philips. 1989. An approximation to the response time for shortest queue routing. *Performance Evaluation Review* **17**(1), 181–189.

Neuts, M.F. 1981. *Matrix-Geometric Solutions in Stochastic Models.* Johns Hopkins University Press, Baltimore.

Zhao, Y. and W.K. Grassmann. 1990. A solution of the shortest queue model with jockeying–

in terms of traffic intensity ρ. *Naval Research Logistics* **37**, 773–787.

Zhao, Y. 1990. *Shortest Queue Models.* Ph.D. Thesis, University of Saskatchewan.

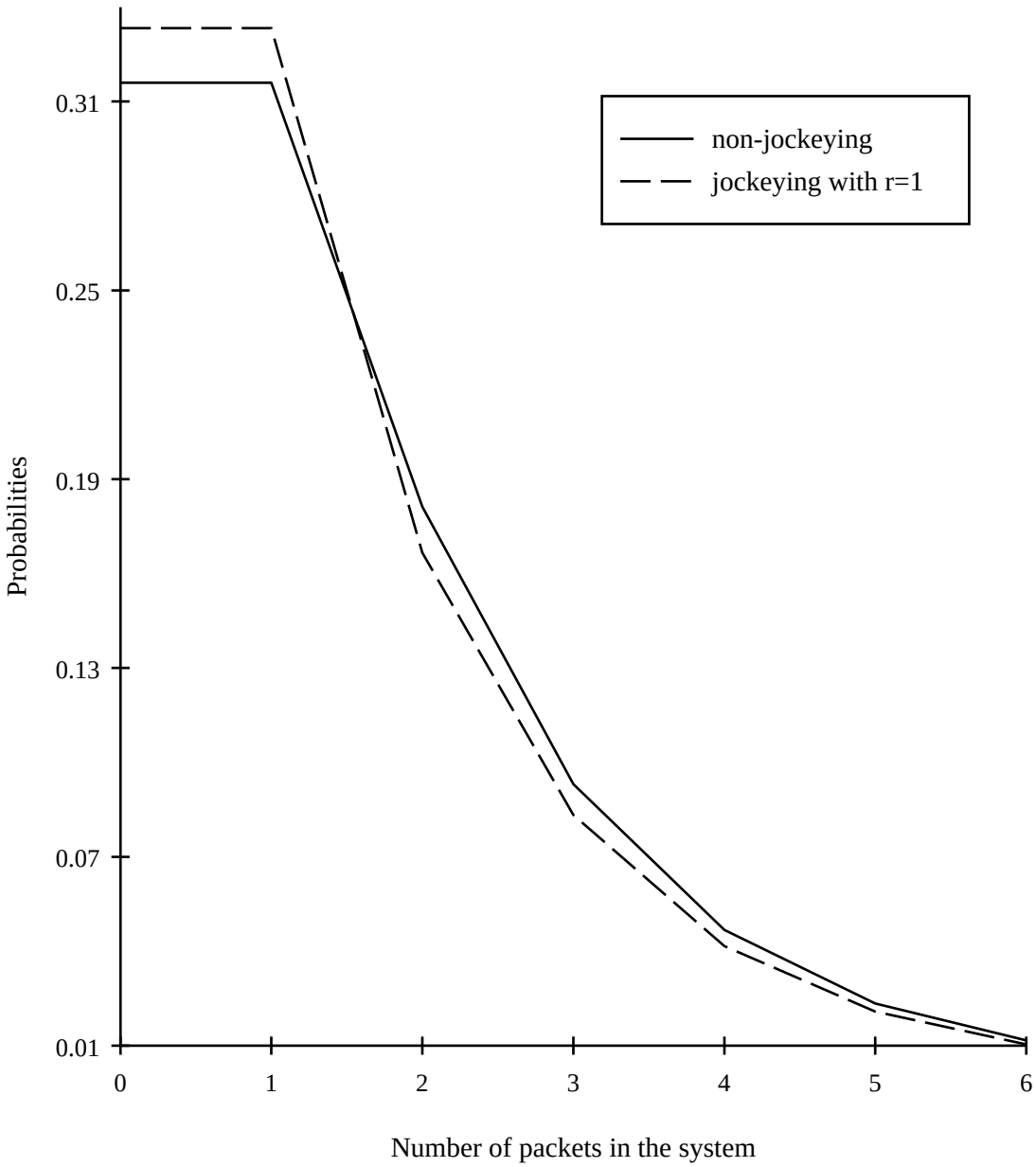

Figure 1: The equilibrium probabilities of the number of packets in the system for the shortest queue model $M/(M/1)^2$ with the traffic intensity $\rho = 0.5$ and equal service rates.

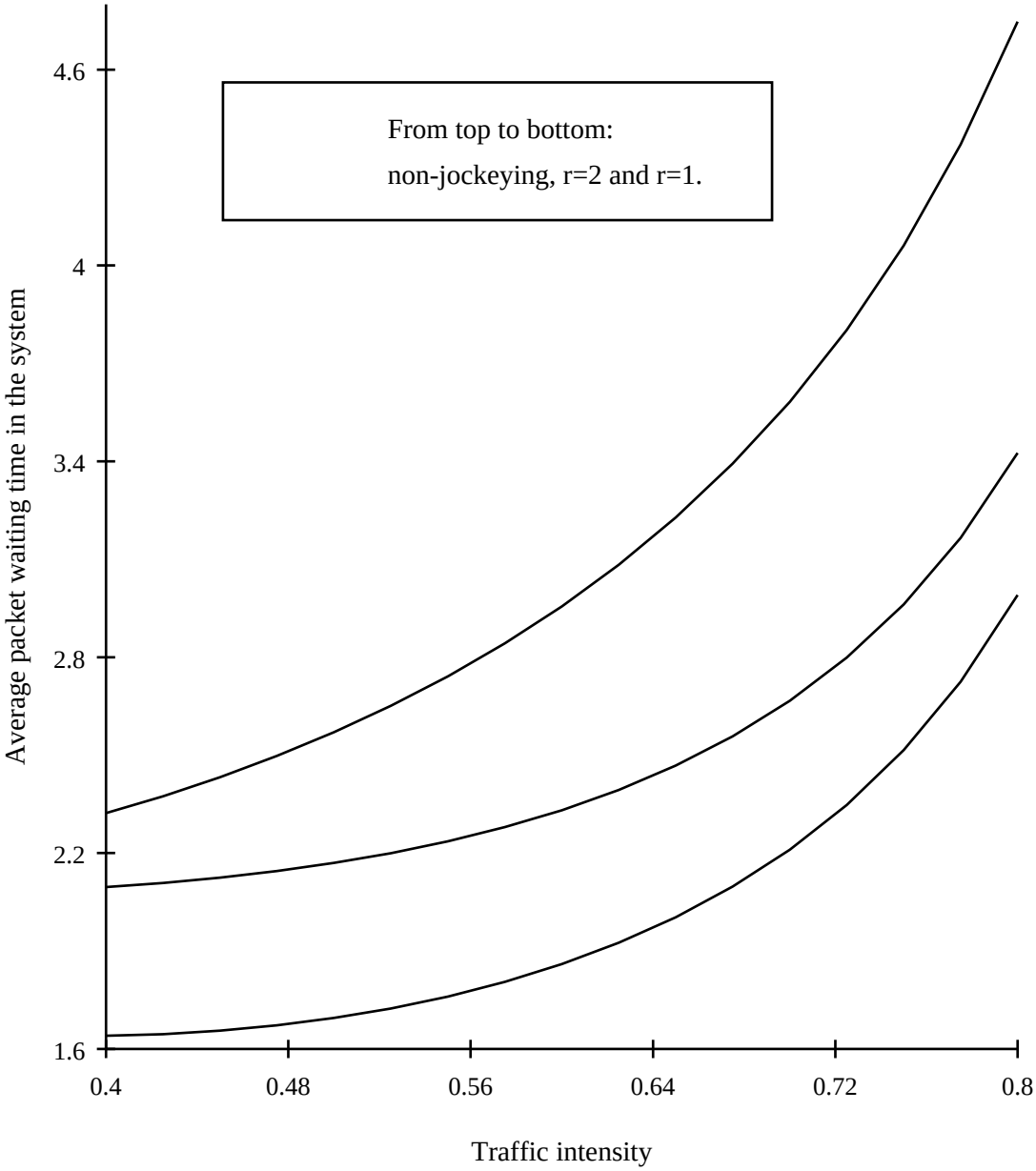

Figure 2: The average packet waiting time in the system for the shortest queue model $M/(M/1)^2$ with the ratio of the service rates 9 to 1.

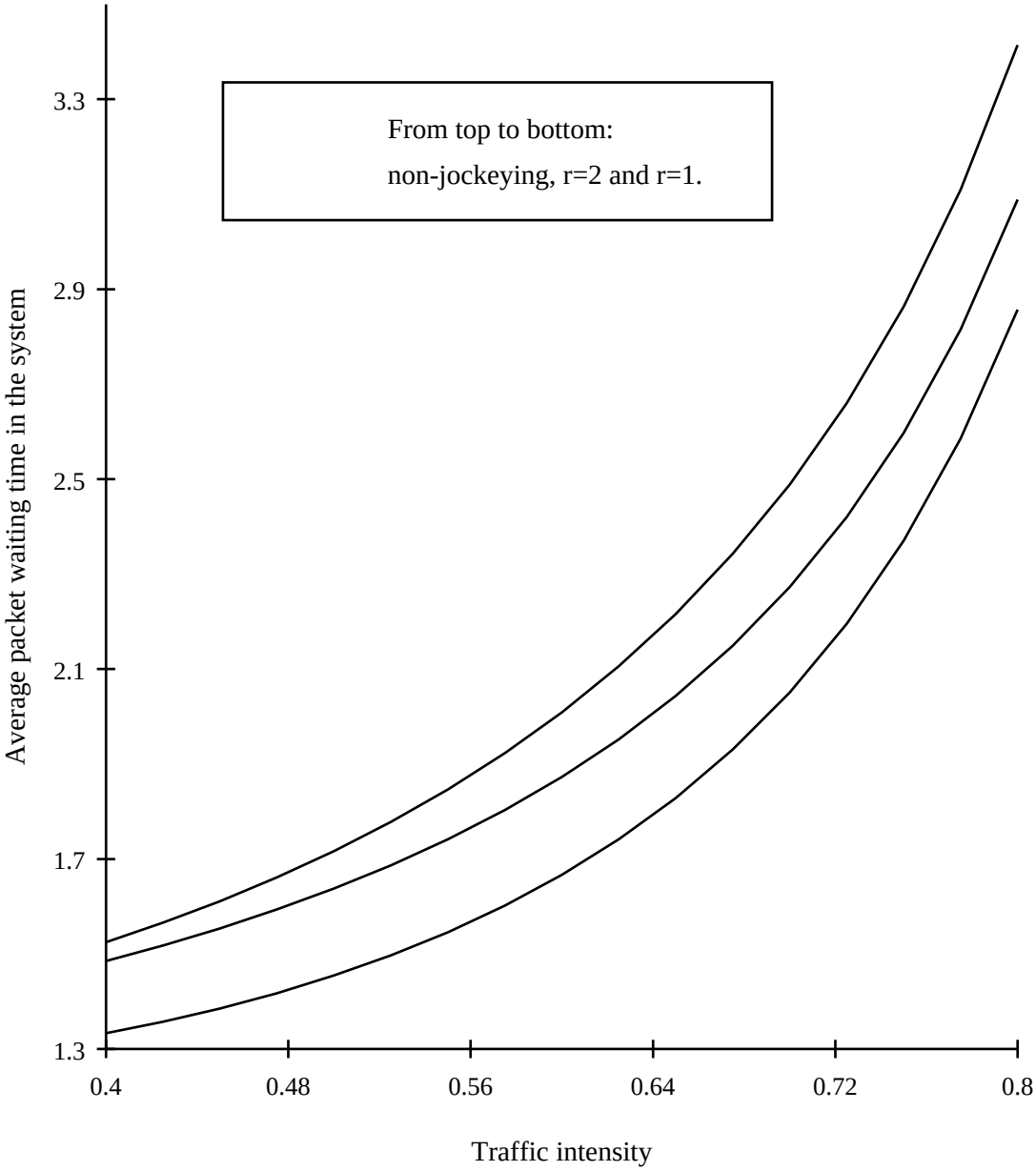

Figure 3: The average packet waiting time in the system for the shortest queue model $M/(M/1)^2$ with the ratio of the service rates 3 to 1.

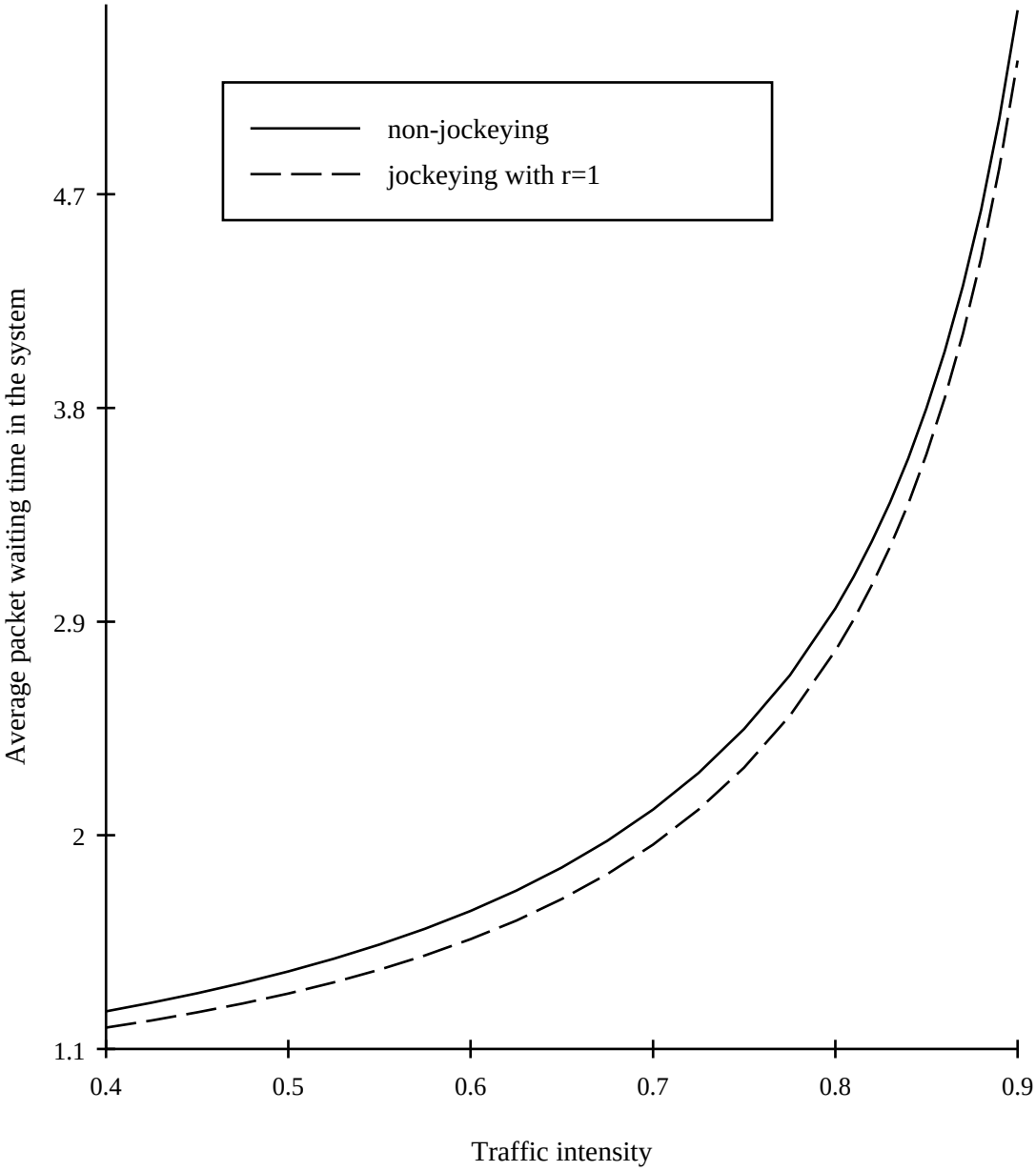

Figure 4: The average packet waiting time in the system for the shortest queue model $M/(M/1)^2$ with the equal service rates.