Journal of the Operations Research
Society of Japan
Vol. 39, No. 1, March 1996

A PREEMPTIVE PRIORITY QUEUE AS A MODEL WITH SERVER VACATIONS

Fumiaki Machihara
NTT Laboratories

(Received April 20, 1994; Revised December 1, 1994)

Abstract We derive the performance measures for non-priority customers in a priority single-server queue with two types of Markovian arrival processes (MAPs). One type of customers has preemptive resume (or repeat) priority over the other. We regard the model with two types of MAP arrivals as an MAP/SM/1 queue with only non-priority customers arrivals, semi-Markovian service times and some server vacations. Performance measures such as the waiting time distribution and the queue length distribution are represented by matrix exponential forms. Our model cannot assume i.i.d. services or i.i.d. server vacations and, in addition, the vacation length depends on the service process. Thus, our formulas give new insights on existing results concerning an MAP/SM/1 queue with i.i.d. multiple server vacations.

1. Introduction

Preemptive priority queues are basic models in queueing theory and have been studied by many researchers [1, 3, 6, 7, 22, 23]. In this paper, we consider the preemptive priority queue in general and unify the existing results.

We consider a single-server queue with two independent non-renewal arrival processes of type 1 customers and type 2 customers. These arrival processes are assumed to be Markovian arrival processes (MAPs) [8] or PHase-type Markov Renewal Processes (PH-MRPs) [11, 12]. One type of customer has preemptive (resume or repeat) priority over the other type of customer. Both types of customers have general service time distributions. Service time distributions for different types of customers may be different. Using an MAP/GI/1 queueing theory, the priority customers' performance measures are analyzed. We focus our attention on the non-priority customers. The services of non-priority customers are preempted and interrupted by priority customers. The time period from the starting point of non-priority customers' service to the end point is influenced by priority customers' arrivals and their service times. We regard this time, sometimes called completion time, as the service time of the non-priority customers. If priority customers' arrival process is not Poissonian, the service process for non-priority customers is not i.i.d. and becomes semi-Markovian. Moreover, a non priority customer may not be able to receive his service because there may be priority customers in the system, even if there are no other non-priority customers at his arrival epoch. Thus, this non-priority customer must wait until the busy period caused by priority customers ends. That is, the non-priority customer encounters a server vacation. Thus, we can consider our model as an MAP/SM/1 queue with semi-Markov services and some server vacations. The server takes a vacation more complex than multiple server vacations [2, 8]. The inter-vacation times are not i.i.d. and depend on the service times, because the busy period length, caused by priority customers, depends on the arrival phase states of priority customers. Moreover, when the server returns from vacation, even if there are no customers in the system, the server spends some time in the system.

In Section 2, we consider the embedded Markov renewal process at non-priority customers' service completion epochs. The transition probability matrix between two successive embedded points and the LST of completion time (service time for non-priority customers) distribution have matrix-exponential forms. We obtain a computable formula for the stationary queue length distribution. We also derive $\mathbf{x}_0$ which is the probability that when a non-priority customer leaves the system, there are no other customers in the system. The fundamental period distribution of the MAP/SM/1 plays an essential role to obtain the $\mathbf{x}_0$. We derive the matrix-exponential form [10, 11, 12, 17, 21, 22] for the LST of this fundamental period distribution. This formula is obtained by using a preemptive LIFO (Last-in-First-out) argument [10, 11, 12, 22] and is represented by double recursion forms.

In section 3, we represent the work-load distribution and waiting time distribution by using the results in section 2. The LIFO preemption argument also plays an important role to derive simple forms for the LSTs of the above distributions. In addition, we derive the relationship between the queue length distribution at the service completion epoch and the waiting time distribution.

2. Model and Analysis

2.1 Model

We analyze a single-server queueing model with two types of customers. Type 1 customers have preemptive resume or repeat priority over type 2 customers. We assume that type $i (i = 1, 2)$ customers arrival process follows a Markovian Arrival Process [8] (MAP) or a PHase-type Markov Renewal Process [11] (PH-MRP) and the interarrival time density has an $n_i \times n_i$ matrix form as

$$f_i(x) = \alpha_i \exp\,(T_i x) T_i^0, \tag{2.1}$$

where α_i, T_i and T_i^0 have the $n_i \times m_i, m_i \times m_i$ and $m_i \times n_i$ matrix forms, respectively.

We may write (2.1) as $f_i(x) = \exp(T_i x) T_i^0 \alpha_i$. In [8], $T_i = C$ and $T_i^0 \alpha_i = D$. In order to emphasize the similarity of MAP or PH-MRP to the phase type renewal process [15], we adopt the notations α_i, T_i and T_i^0 [15, 16]. When α_i and T_i^0 are a row vector and a column vector, respectively, MAP becomes a phase type renewal process. When $T_i^0 \alpha_i$ is diagonal, MAP becomes a Markov modulated Poisson process [4, 5].

The service time distributions of type 1 customers and type 2 customers are given by $H_1(x)$ and $H_2(x)$, respectively.

Since the arrival and service processes for type 1 customers are not influenced by those of type 2 customers, the performance measures for type 1 customers can be analyzed by the existing MAP/GI/1 theory [9, 18, 24].

Let us study the performance measures of type 2 customers. We consider the embedded Markov renewal process of the service completion epochs of type 2 customers. Let $z_1, z_2, \cdots$ denote successive service completion epochs of type 2 customers. Let $N_k^*, J_k^*(1)$ and $J_k^*(2)$ denote the number of type 2 customers in system, the arrival phase state of type 1 customers and the arrival phase state of type 2 customers at epoch $z_k + 0$ immediately after $z_k (k = 1, 2, \cdots)$, respectively. Throughout this section, we consider a stable queue, that is,

$$\rho = \rho_1 + \rho_2 < 1, \tag{2.2}$$

where ρ_i is the server utilization by type i customers, $i = 1, 2$. For the mean service time μ_i^{-1} of type i customers, ρ_i is given by

$$\rho_i = (\mathbf{q}_i \alpha_i (-T_i)^{-2} T_i^0 \mathbf{e})^{-1} \mu_i^{-1} = \lambda_i \mu_i^{-1}, \tag{2.3}$$

where $\mathbf{e}$ is a column vector with all elements equal to 1, and $\mathbf{q}_i$ is the invariant probability vector for $\alpha_i (-T_i)^{-1} T_i^0$.

2.2 Stationary queue length distribution

If our queue is stable, we can define the stationary probability vector and the associated generating function as follows:

$$\begin{aligned}\mathbf{x}_i =&(x_i(1,1),\cdots,x_i(1,m_2),x_i(2,1),\cdots,x_i(2,m_2),\cdots,x_i(m_1,1),\cdots,\\ &x_i(m_1,m_2),\end{aligned} \tag{2.4}$$

where

$$x_i(j,k) = \lim_{n\to\infty} P\{N_n^* = i, J_n^*(1) = j, J_n^*(2) = k\} \tag{2.5}$$

and

$$\mathrm{g}(z) = \sum_{j=0}^{\infty} z^i \mathbf{x}_i, \qquad | z |\leq 1. \tag{2.6}$$

Hereafter, states (i,j), where i and j are the arrival phase states of type 1 customers and type 2 customers, respectively, are always in lexicographic order, that is, $(1,1)$, $\cdots$, $(1,m_2)$, $(2,1)$, $\cdots$, $(2,m_2)$, $\cdots$, $(m_1,1)$, $\cdots$, (m_1,m_2).

Now, we consider the transition probability matrix P between two successively embedded Markov renewal points. P can be written as follows:

$$P = \begin{bmatrix} B_0 & B_1 & B_2 & B_3 & \dots \\ A_0 & A_1 & A_2 & A_3 & \dots \\ 0 & A_0 & A_1 & A_2 & \dots \\ 0 & 0 & A_0 & A_1 & \dots \\ 0 & 0 & 0 & A_0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}, \tag{2.7}$$

where (i,j) element is the transition probability matrix from the number of customer i to j. Each element is given by

$$A_i = \int_0^{\infty} (dC(x)\otimes I_2)(I_1 \otimes P_i^{(2)}(x)), \qquad i = 0,1,\cdots \tag{2.8}$$

and

$$\begin{aligned} B_{i-1} =& [-T - T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(-I_1\otimes T_2)]^{-1} \\ &\cdot [(T_1^0\alpha_1 \otimes I_2)\{\int_0^{\infty}(dG_1(\cdot) * C(x)\otimes I_2)(I_1\otimes P_i^{(2)}(x)) \\ &- \int_0^{\infty}(dG_1(x)\otimes I_2)(I_1\otimes P_0^{(2)}(x))\cdot \int_0^{\infty}(dC(x)\otimes I_2)(I_1\otimes P_i^{(2)}(x))\} \\ &+ (I_1\otimes T_2^0\alpha_2)\int_0^{\infty}(dC(x)\otimes I_2)(I_1\otimes P_i^{(2)}(x))],\ i = 1,2,\cdots. \end{aligned} \tag{2.9}$$

Here, I_j is the $m_j \times m_j$ identity matrix $(j = 1,2), T = T_1 \otimes I_2 + I_1 \otimes T_2$, and $P_i^{(2)}(x)(i = 0,1,\cdots)$ is the probability matrix that the number of arrivals of non-priority customers in $[0,x]$ is exactly i. The distribution matrix $G_1(x)$ is the fundamental period length distribution of type 1 (priority) customers' queueing process; its LST is given by [11, 12, 17, 19]

$$\begin{aligned} G_1^*(s) &= H_1^*(sI_1 - T_1 - T_1^0\alpha_1 G_1^*(s)) \\ &= \int_0^{\infty} dH_1(x)\exp\{-(sI_1 - T_1 - T_1^0\alpha_1 G_1^*(s))x\}. \end{aligned} \tag{2.10}$$

In addition,

$$G_1^*(-I_1 \otimes T_2) = \int_0^\infty (dG_1(x) \otimes I_2)\exp\{(I_1 \otimes T_2)x\} \tag{2.11}$$

and

$$dG_1(\cdot) * C(x) = \int_0^x dG_1(y)C(x-y). \tag{2.12}$$

An $m_1 \times m_1$ matrix $C(x)$ is the completion time distribution, and is defined as the duration of the period that simultaneously begins from the instant a non-priority customer's service starts and ends at the instant the server becomes free to serve the next non-priority customer. The LST $C^*(s)$ ($Re(s) \geq 0$) of $C(x)$ is given for each preemptive discipline as follows [13]:

$$\begin{aligned} C^*(s) &= H_2^*(sI_1 - T_1 - T_1^0\alpha_1 G_1^*(s)) \\ &= \int_0^\infty dH_2(x)\exp\{-(sI_1 - T_1 - T_1^0\alpha_1 G_1^*(s))x\} \end{aligned} \tag{2.13}$$

for the Preemptive Resume Discipline (PRS);

$$\begin{aligned} C^*(s) = \int_0^\infty dH_2(x)[I_1 - \{I_1 - \exp(-x(sI_1 - T_1))\}(sI_1 - T_1)^{-1}T_1^0\alpha_1 G_1^*(s)]^{-1} \\ \cdot \exp(-x(sI_1 - T_1)) \end{aligned} \tag{2.14}$$

for the Preemptive Repeat-Identical Discipline (PRI);

$$C^*(s) = [I_1 - \{I_1 - H_2^*(sI_1 - T_1)\}(sI_1 - T_1)^{-1}T_1^0\alpha_1 G_1^*(s)]^{-1} H_2^*(sI_1 - T_1), \tag{2.15}$$

where

$$H_2^*(sI_1 - T_1) = \int_0^\infty dH_2(x)\exp\{-(sI_1 - T_1)x\},$$

for the Preemptive Repeat-Different Discipline (PRD).

Note that $G_1^*(-I_1 \otimes T_2)$ in (2.9) can be obtained by successive substitution as follows:

$$\begin{aligned} &G_1^{*(0)}(-I_1 \otimes T_2) = H_1^*(-I_1 \otimes T_2 - T_1 \otimes I_2), \\ &G_1^{*(n+1)}(-I_1 \otimes T_2) = H_1^*(-I_1 \otimes T_2 - T_1 \otimes I_2 - T_1^0\alpha_1 \otimes I_2 \cdot G_1^{*(n)}(-I_1 \otimes T_2)), \\ &\qquad n = 0, 1, \cdots. \end{aligned} \tag{2.17}$$

If we consider the non-priority customers' queueing process, we can regard a preemptive priority model with two types of MAP inputs as a model with single MAP inputs, semi-Markov services and some server vacations, that is, as an $MAP_2/SM/1$ with some server vacations. We may consider the busy period caused by priority customers which is started by a priority customer arriving when the number of customers in the system is equal to 0, as a server vacation period. The transition probability matrix B_{i-1} in (2.9) is obtained by the following consideration. The term

$$\begin{aligned} &[-T - T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(-I_1 \otimes T_2)]^{-1}[T_1^0\alpha_1 \otimes I_2\{\int_0^\infty (dG(\cdot) * C(x) \otimes I_2)(I_1 \otimes P_i^{(2)}(x)) \\ &- \int_0^\infty (dG_1(x) \otimes I_2)(I_1 \otimes P_0^{(2)}(x)) \cdot \int_0^\infty (dC(x) \otimes I_2)(I_1 \otimes P_i^{(2)}(x))\}] \end{aligned}$$

corresponds to the case in which i non-priority customers arrive during the server vacation, that is, the busy period caused by priority customers. The term

$$[-T - T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(-I_1 \otimes T_2)]^{-1}[I_1 \otimes T_2^0\alpha_2 \int_0^\infty (dC(x) \otimes I_2)(I_1 \otimes P_i^{(2)}(x))]$$

corresponds to the case in which a non-priority customer simultaneously receives service at its arrival and i other non-priority customers arrive during its service time (completion time). The transition probability matrix A_i can be easily obtained, when we assume that the service time distribution is $C(x)$.

Theorem 2.1

The generating function $\mathbf{g}(z)$ for $|z| \leq 1$ is given by

$$\begin{aligned}\mathbf{g}(z) &= \mathbf{x}_0[-T - T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(-I_1 \otimes T_2)]^{-1} \\ &\cdot [I_1 \otimes T_2^0\alpha_2 z + T + T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z))] \\ &\cdot C^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z))(zI - C^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z)))^{-1}, \end{aligned} \tag{2.18}$$

where

$$C^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z)) = \int_0^\infty (dC(x) \otimes I_2)(I_1 \otimes \exp(T_2 + T_2^0\alpha_2 z)x). \tag{2.19}$$

Proof

From the definition of a stationary distribution, we have

$$(\mathbf{x}_0, \mathbf{x}_1, \cdots) = (\mathbf{x}_0, \mathbf{x}_1, \cdots)P. \tag{2.20}$$

Hence, when $A(z) = \sum_{i=0}^{\infty} A_i z^i$ and $B(z) = \sum_{i=0}^{\infty} B_i z^i$, we obtain

$$\begin{aligned}\mathbf{g}(z) &= \mathbf{x}_0 B(z) + \Big(\frac{1}{z}\sum_{i=1}^{\infty} x_i z^i\Big)A(z) \\ &= \mathbf{x}_0 B(z) + \frac{1}{z}(\mathbf{g}(z) - \mathbf{x}_0)A(z), \end{aligned} \tag{2.21}$$

thus

$$\mathbf{g}(z)(zI - A(z)) = \mathbf{x}_0(zB(z) - A(z)). \tag{2.22}$$

Since

$$\sum_{i=0}^{\infty} z^i P_i^{(2)}(x) = \exp\{(T_2 + T_2^0\alpha_2 z)x\},$$

we have from (2.8)

$$A(z) = C^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z)). \tag{2.23}$$

Since

$$\begin{aligned}&\int_0^\infty (dG_1(\cdot) * C(x) \otimes I_2)(I_1 \otimes P_0^{(2)}(x)) \\ &= \int_0^\infty (dG_1(x) \otimes I_2)(I_1 \otimes P_0^{(2)}(x)) \cdot \int_0^\infty (dC(x) \otimes I_2)(I_1 \otimes P_0^{(2)}(x)),\end{aligned}$$

we have from (2.9)

$$\begin{aligned}B(z) &= [-T - T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(-I_1 \otimes T_2)]^{-1} \\ &\cdot [I_1 \otimes T_2^0\alpha_2 + \frac{1}{z}(T_1^0\alpha_1 \otimes I_2)(G_1^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z)) - G_1^*(-I_1 \otimes T_2))] \\ &\cdot C^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z)). \end{aligned} \tag{2.24}$$

Substituting (2.23) and (2.24) into (2.22), we have (2.18). □

Remark It is an important assumption on computation that both H_1 and H_2 have scalar structures. Both $G_1(\cdot)$ for (2.10) and $C(\cdot)$ for (2.13) or (2.15) have a common matrix-exponential form as follows:

$$\int_0^\infty dH(x)\exp(-Ax),$$

where $H(x)$ is scalar and A is a matrix. This value can be computed by substituting "A" into "s" in the LST $H^*(s)$ of $H(x)$. When H has a matrix or vector structure such as the case in semi-Markov services, we cannot simply substitute "A" into "s" in $H^*(s)$. For example, if H_1 has an $n\times n$ matrix structure in (2.10) such as semi-Markov services, we have to compute a matrix-exponential form

$$G_1^*(s) = \int_0^\infty dH_1(x)\otimes I_1\cdot\exp\{-[s(I_n\otimes I_1) - I_n\otimes T_1 - (I_n\otimes T_1^0\alpha_1)(G_1^*(s)]x\},$$

where I_n is an $n\times n$ identity matrix.

2.3 How to obtain $\mathbf{x}_0$

In order to obtain the unknown probability $\mathbf{x}_0$ in Theorem 2.1, we consider the fundamental period of $MAP_2/SM/1$ with or without server vacations. The fundamental period is defined as the first-passage time from i (the number of customers i in the system) to $i-1$ for $i=1,2,\cdots$. Note that the fundamental period length is independent of whether or not the server has vacations, and is also independent of i. We define

$$\begin{aligned} B(t) &= (B_{i_1,i_2}^{(j_1j_2)}(t); i_1,i_2=1,2,\cdots,m_1, j_1,j_2=1,2,\cdots,m_2) \\ &= (Pr\{T_f\le t, J_{T_f}^*(1)=i_2, J_{T_f}^*(2)=j_2 \mid J_0^*(1)=i_1, J_0^*(2)=j_1\}) \end{aligned} \tag{2.25}$$

for the fundamental period length T_f. In this definition, we interpret $J_t^*(i)$ as the arrival phase state of type i customers at time t.
The LST $B^*(s)$ of $B(t)$ has the form

$$\begin{aligned} B^*(s) &= \sum_{n=0}^\infty\{\int_0^\infty e^{-sx}(dC(x)\otimes I_2)(I_1\otimes P_n^{(2)}(x))\}B^{*n}(s) \\ &= \int_0^\infty (dC(x)\otimes I_2)\exp\{(-sI+I_1\otimes T_2+I_1\otimes T_2^0\alpha_2\cdot B^*(s))x\} \\ &= C^*(sI-I_1\otimes T_2-I_1\otimes T_2^0\alpha_2\cdot B^*(s)). \end{aligned} \tag{2.26}$$

The third term can be obtained by using the LIFO (Last-In-First-Out) preemption argument [10, 11, 12]. For example, in the proof of Theorem 3.1 in [12], we may assume that $T=I_1\otimes T_2$, $T^0\alpha = I_1\otimes T_2^0\alpha_2$ and $G^*(s,1-)=B^*(s)$. Since the service time distribution is $C(x)\otimes I_2$, we obtain the third term. Because of the forms of $C^*(s)$, (2.13) through (2.15), the fourth term can be written. The fourth term indicates that we may simply substitute matrix $sI-I_1\otimes T_2-I_1\otimes T_2^0\alpha_2\cdot B^*(s)$ into s in (2.13) (or (2.14) or (2.15)). For example, when the service discipline is PRS, we have

$$\begin{aligned} B^*(s) &= H_2(sI-I_1\otimes T_2-I_1\otimes T_2^0\alpha_2\cdot B^*(s)-T_1\otimes I_2 \\ &\quad - T_1^0\alpha_1\otimes I_2\cdot G_1^*(sI-I_1\otimes T_2-I_1\otimes T_2^0\alpha_2\cdot B(s))). \end{aligned}$$

When the service discipline is PRI (or PRD), we may use (2.14) (or (2.15)).

Let us consider computation of the matrix $B^*(s)$ for $Re(s) \geq 0$. This matrix may be computed by the following double recursive forms. In particular, we consider the case in which the service discipline is PRS.

At first, compute $G_1^*(W_0)$ for $W_0 = sI - I_1 \otimes T_2$ by successive substitutions in (2.10), that is,

$$G_1^{*(0)}(W_0) = 0,$$
$$G_1^{*(m+1)}(W_0) = H_1^*(W_0 - T_1 \otimes I_2 - T_1^0\alpha_1 \otimes I_2 \cdot G_1^{*(m)}(W_0)), \qquad m = 0, 1, \cdots.$$

Next, compute $B^*(s)$, starting with

$$B^{*(0)}(s) = C^*(W_0) = H_2^*(W_0 - T_1 \otimes I_2 - T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(W_0)) \tag{2.27}$$

and executing the following steps until $B^{(n)*}(s)$ converges to $B^*(s)$ in (2.13).

(Step 1) Compute $G_1^*(W_n)$ for $W_n = sI - I_1 \otimes T_2 - I_1 \otimes T_2^0\alpha_2 \cdot B^{*(n-1)}(s)$ by successive substitutions, that is,

$$G_1^{*(0)}(W_n) = 0,$$
$$G_1^{*(m+1)}(W_n) = H_1^*(W_n - T_1 \otimes I_2 - T_1^0\alpha_1 \otimes I_2 \cdot G_1^{*(m)}(W_n)), \qquad m = 0, 1, \cdots.$$

(Step 2) Compute $B^{*(n)}(s)$ by

$$B^{*(n)}(s) = C^*(W_n) = H_2^*(W_n - T_1 \otimes I_2 - T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(W_n)), \qquad n = 1, 2, \cdots. \tag{2.28}$$

When the service discipline is PRI, we may use the following formulas instead of Equations (2.27) and (2.28).

$$C^*(W_0) = \int_0^\infty dH_2(x)[I - \{I - \exp(-x(W_0 - T_1 \otimes I_2))\}(W_0 - T_1 \otimes I_2)^{-1} \cdot T_1^0\alpha_1 G_1^*(W_0)]^{-1} \cdot \exp(-x(W_0 - T_1 \otimes I_2)) \tag{2.27A}$$

and

$$C^*(W_n) = \int_0^\infty dH_2(x)[I - \{I - \exp(-x(W_n - T_1 \otimes I_2))\}(W_0 - T_1 \otimes I_2)^{-1} \cdot T_1^0\alpha_1 \otimes G_1^*(W_n)]^{-1} \cdot \exp(-x(W_n - T_1 \otimes I_2)). \tag{2.28A}$$

When the service discipline is PRD, we may use the following formulas.

$$C^*(W_0) = [I - \{I - H_2^*(W_0 - T_1 \otimes I_2)\}(W_0 - T_1 \otimes I_2)^{-1}T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(W_0)]^{-1} \cdot H_2^*(W_0) - T_1 \otimes I_2) \tag{2.27B}$$

and

$$C^*(W_n) = [I - \{I - H_2^*(W_n - T_1 \otimes I_2)\}(W_n - T_1 \otimes I_2)^{-1}T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(W_n)]^{-1} \cdot H_2^*(W_n - T_1 \otimes I_2). \tag{2.28B}$$

We can easily prove for $s \geq 0$ that $B^{*(n+1)}(W_n)$ is greater than or equal to $B^{*(n)}(W_{n-1})$ elementwise. This means $B^{*(n)}(W_{n-1})$ converges to the minimal nonnegative solution of (2.26) as $n \to \infty$.

Now, we consider successive returns to $N_k^* = 0$ for the embedded Markov renewal process. The LST of the successive return interval distribution is given by

$$A(s) = \sum_{n=0}^{\infty} \{(sI - T)^{-1} T_1^0 \alpha_1 \otimes I_2 \cdot G_1^*(sI - I_1 \otimes T_2)\}^n (sI - T)^{-1}$$
$$\cdot [T_1^0 \alpha_1 \otimes I_2 \sum_{j=1}^{\infty} \int_0^{\infty} e^{-sx} dG_1(x) \otimes I_2 (I_1 \otimes P_j^{(2)}(x)) \cdot B^{*j}(s) + I_1 \otimes T_2 \alpha_2 \cdot B^*(s)]$$
$$= [sI - T - T_1^0 \alpha_1 \otimes I_2 \cdot G_1^*(sI - I_1 \otimes T_2)]^{-1}$$
$$\cdot [T_1^0 \alpha_1 \otimes I_2 \{G_1^*(sI - I_1 \otimes T_2 - I_1 \otimes T_2^0 \alpha_2 \cdot B^*(s)) - G_1^*(sI - I_1 \otimes T_2)\}$$
$$+ I_1 \otimes T_2^0 \alpha_2 \cdot B^*(s)]. \qquad (2.29)$$

Theorem 2.3

The vector $\mathbf{x}_0$ is given for the arrival rate λ_2 of type 2 customers by

$$\mathbf{x}_0 = \mathbf{x}_0[-T - T_1^0 \alpha_1 \otimes I_2 \cdot G_1^*(-I_1 \otimes T_2)]^{-1}$$
$$\cdot [T_1^0 \alpha_1 \otimes I_2 \{G_1^*(-I_1 \otimes T_2 - I_1 \otimes T_2^0 \alpha_2 \cdot B^*(0)) - G_1^*(-I_1 \otimes T_2)\} + I_1 \otimes T_2^0 \alpha_2 \cdot B^*(0)] \qquad (2.30)$$

and

$$\lambda_2 \mathbf{x}_0[-T - T_1^0 \alpha_1 \otimes I_2 \cdot G_1^*(-I_1 \otimes T_2)]^{-1} \mathbf{e} = 1 - \rho, \qquad (2.31)$$

where $\mathbf{e}$ is a column vector whose elements are equal to one.

Proof

Equation (2.30) is the definition of $\mathbf{x}_0$ itself.

Let $\mathbf{p}_0 = (p_0(1,1), \cdots, p_0(1, m_2), \cdots, p_0(m_1, 1), \cdots, p_0(m_1, m_2))$ denote the probability vector that the server is idle at an arbitrary time. Let $M_{j1,j2}(0;t)$ denote the expected number of visits, by the embedded Markov renewal process of the non-priority customer service completion epochs, to the state $(0, j_1, j_2), j_1 = 1, \cdots, m_1, j_2 = 1, \cdots, m_2$, in $[0, t]$. We do not need the explicit form of the Markov renewal matrix $M_{j_1,j_2}(0;t)$, because we need the form $M_{j_1,j_2}(0;t)$ only for sufficiently large t in the following discussion and $M_{j_1,j_2}(0;t)$ does not depend on the initial value. Here, j_i is the arrival phase state of type i customers, $i = 1, 2$. Then,

$$\mathbf{p}_0 = \lim_{t \to \infty} \int_0^t dM(0;u) \sum_{n=0}^{\infty} [(-T)^{-1}(T_1^0 \alpha_1 \otimes I_2) G_1^*(-I_1 \otimes T_2)]^n \exp(T(t-u)), \qquad (2.32)$$

where

$$\mathbf{M}(0,u) = (M_{1,1}(0,u), \cdots, M_{1,m_2}(0,u), \cdots, M_{m_1,1}(0,u), \cdots, M_{m_1,m_2}(0,u)).$$

Using the key renewal theory, we obtain

$$\mathbf{p}_0 = \lambda_2 \mathbf{x}_0 \sum_{n=0}^{\infty} [(-T)^{-1}(T_1^0 \alpha_1 \otimes I_2) G_1^*(-I_1 \otimes T_2)]^n (-T)^{-1}$$
$$= \lambda_2 \mathbf{x}_0[-T - (T_1^0 \alpha_1 \otimes I_2) G_1^*(-I_1 \otimes T_2)]^{-1}. \qquad (2.33)$$

Since

$$\mathbf{p}_0 \mathbf{e} = 1 - \rho, \qquad (2.34)$$

we obtain (2.31). □

3. The stationary waiting time distribution

In this section, we study the joint probability of the virtual waiting time of non-priority customers and the arrival phase states of priority and non-priority customers. The $V_{N,j_1,j_2}(x)$, $x \geq 0$, $j_1 = 1, \cdots, m_1$, $j_2 = 1, \cdots, m_2$, denotes the stationary probability that a virtual customer arriving at arbitrary time has to wait at most x and finds the arrival processes of priority and non-priority customers in the phases, j_1 and j_2, respectively. When we define the vector distribution

$$\mathbf{V_N}(x) = (V_{N,1,1}(x), \cdots, V_{N,1,m_2}(x), \cdots, V_{N,m_1,1}(x), \cdots, V_{N,m_1,m_2}(x)), \tag{3.1}$$

we have the following theorem.

Theorem 3.1

For a stable queue, the LST of $\mathbf{V_N}(x)$ satisfies the equation

$$\begin{aligned}&\mathbf{V}^*_\mathbf{N}(s)(sI + I_1 \otimes T_2 + I_1 \otimes T_2^0\alpha_2 \cdot C^*(s) \otimes I_2)\\ &= \lambda_2\mathbf{x}_0[-T - T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(-I_1 \otimes T_2)]^{-1}(sI - T_1 \otimes I_2 - T_1^0\alpha_1 G_1^*(s) \otimes I_2)\\ &= \mathbf{p}_0(sI - T_1 \otimes I_2 - T_1^0\alpha_1 G_1^*(s) \otimes I_2).\end{aligned} \tag{3.2}$$

Proof

First, we introduce the corresponding time dependent distribution $\mathbf{V_N}(t;x)$ at time t. Let $M_{j_1,j_2}(i;t)$ denote the expected number of visits by the embedded Markov renewal process associated with the non-priority customer service completion epochs to the state $(i,j_1,j_2), i \geq 0, j_1 = 1, \cdots, m_1, j_2 = 1, \cdots, m_2$, in $[0,t]$. We do not need the explicit form of the Markov renewal matrix $M_{j_1,j_2}(i;t)$, because we need the form $M_{j_1,j_2}(i;t)$ only for sufficiently large t in the following discussion and $M_{j_1,j_2}(i;t)$ does not depend on the initial value. When we define the vector

$$\mathbf{M}(i;t) = (M_{11}(i;t), \cdots, M_{1,m_2}(i;t), \cdots, M_{m_1,1}(i;t), \cdots, M_{m_1,m_2}(i;t)),$$

we have the following equation.

$$\begin{aligned}\mathbf{V_N}(t;x) &= \int_0^t d\mathbf{M}(0;\cdot) * N(u)\exp(T(t-u))\\ &+ \int_{0_{(u)}}^t \int_{0_{(v)}}^{t-u} \int_{0_{(w)}}^x d\mathbf{M}(0;\cdot) * N(u)\exp(Tv)\\ &\cdot [\sum_{i=0}^{\infty}(T_1^0\alpha_1 \otimes I_2)dv\{(I_1 \otimes P_i^{(2)}(t-u-v))(dG_1(\cdot) * C(t+w-u-v) \otimes I_2)(C^{[i-1]}(x-w) \otimes I_2)\\ &- \int_{0_{(y)}}^{\infty}(I_1 \otimes P_0^{(2)}(y))(dG_1(y) \otimes I_2)(I_1 \otimes P_i^{(2)}(t-u-v))(dC(t+w-u-v) \otimes I_2)\\ &\cdot C^{[i-1]}(x-w) \otimes I_2\}]\\ &+ \int_{0_{(u)}}^t \int_{0_{(v)}}^{t-u} \int_{0_{(w)}}^x d\mathbf{M}(0;\cdot) * N(u)\exp(Tv)\sum_{i=0}^{\infty}(I_1 \otimes T_2^0\alpha_2)dv(I_1 \otimes P_i^{(2)}(t-u-v))\\ &\cdot (dC(t+w-u-v) \otimes I_2)(C^{[i]}(x-w)) \otimes I_2)\\ &+ \sum_{i=1}^{\infty}\sum_{k=1}^{i} \int_{0_{(u)}}^t \int_{0_{(w)}}^x d\mathbf{M}(k;u)(I_1 \otimes P_{i-k}^{(2)}(t-u))d(C(t+w-u) \otimes I_2)(C^{[i-1]}(x-w) \otimes I_2),\end{aligned} \tag{3.3}$$

where $P_i^{(2)}(t)$ is the matrix which represents the probability that the number of non-priority customers arrivals in $[0, t]$ is i,

$$dX(\cdot) * Y(x) = \int_0^\infty dX(y)Y(x-y),$$

and $X^{[n]}(x)$ is the n-th convolution of $X(x)$. $N(u)$ is given by

$$\gamma(u) = \int_0^u \exp\ \{T(u-u_1)\}(T_1^0\alpha_1 \otimes I_2)(dG_1(u_1) \otimes I_2)\ \exp\ \{(I_1 \otimes T_2)u_1\}$$

and

$$N(u) = \sum_{n=0}^\infty \gamma^{[n]}(u).$$

The first, second and third terms correspond to the case where the system was empty at the last departure of a non-priority customer prior to t. The first term of the right-hand-side corresponds to the case where no non-priority customers arrive until t, and the system is empty at t. The second term corresponds to the case where the first non-priority customer after 0 arrives during the busy period of priority customers. The third term corresponds to the case where the first non-priority customer finds the system empty. The fourth term is the case where the last non-priority customer served prior to t left k non-priority customers in the system.

Now, we introduce the following symbols to simplify the expressions.

$$Z = -T - (T_1^0\alpha_1 \otimes I_2)G_1^*(-I_1 \otimes T_2), \tag{3.4}$$

$$D_1 = T_1^0\alpha_1 \otimes I_2, \tag{3.5}$$

$$D_2 = I_1 \otimes T_2^0\alpha_2, \tag{3.6}$$

$$q_i(t) = I_1 \otimes P_i^{(2)}(t), \tag{3.7}$$

$$H(t) = (G_1(\cdot) * C(t)) \otimes I_2), \tag{3.8}$$

$$A(t) = C(t) \otimes I_2 \tag{3.9}$$

and

$$G(t) = G_1(t) \otimes I_2. \tag{3.10}$$

Taking the limit as $t \to \infty$ in (3.3), we have from the key renewal theory the following equation. Here, λ_e is the output rate of the non-priority customers and is equal to the input rate λ_2.

$$\begin{aligned}
\mathbf{V_N}(x) &= \lambda_e \mathbf{x}_0 Z^{-1}(-T)[\int_0^\infty \exp\ (Ty)dy \\
&+ \int_{0_{(t)}}^\infty dt \int_{0_{(v)}}^t \int_{0_{(w)}}^x \exp\ (Tv)D_1 dv \sum_{i=0}^\infty q_i(t-v)dH(t+w-v)A^{[i-1]}(x-w) \\
&- \int_{0_{(t)}}^\infty dt \int_{0_{(v)}}^t \int_{0_{(w)}}^x \exp\ (Tv)D_1 dv \int_0^\infty q_0(u)dG(u) \cdot \sum_{i=0}^\infty q_i(t-v)(dA(t+w-v))A^{[i-1]}(x-w) \\
&+ \int_{0_{(t)}}^\infty dt \int_{0_{(v)}}^t \int_{0_{(w)}}^x \exp\ (Tv)D_2 dv \sum_{i=0}^\infty q_i(t-v)dA(t+w-v)A^{[i]}(x-w)]
\end{aligned}$$

$$+\sum_{i=1}^{\infty}\sum_{k=1}^{i}\int_{0_{(t)}}^{\infty}dt\int_{0_{(w)}}^{x}\lambda_e\mathbf{x}_k q_{i-k}(t)dA(t+w)A^{[i-1]}(x-w)$$
$$=\lambda_e\mathbf{x}_0Z^{-1}+\lambda_e\mathbf{x}_0Z^{-1}[D_1\int_{0_{(t)}}^{\infty}dt\int_{0_{(w)}}^{x}\sum_{i=0}^{\infty}q_i(t)dH(t+w)A^{[i-1]}(x-w)$$
$$-D_1\int_{0_{(u)}}^{\infty}q_0(u)dG(u)\int_{0_{(t)}}^{\infty}dt\int_{0_{(w)}}^{x}\sum_{i=0}^{\infty}q_i(t)dA(t+w)A^{[i-1]}(x-w)$$
$$+D_2\int_{0_{(t)}}^{\infty}\int_{0_{(w)}}^{x}\sum_{i=0}^{\infty}q_i(t)dA(t+w)A^{[i]}(x-w)]$$
$$+\sum_{i=1}^{\infty}\sum_{k=1}^{i}\int_{0_{(t)}}^{\infty}dt\int_{0_{(w)}}^{x}\lambda_e\mathbf{x}_k q_{i-k}(t)dA(t+w)A^{[i-1]}(x-w). \quad (3.11)$$

Taking the Laplace-Stieltjes transform of $\mathbf{V_N}(x)$, we obtain

$$\mathbf{V}_{\mathbf{N}}^{*}(s)=\lambda_e\mathbf{x}_0Z^{-1}$$
$$+\lambda_e\mathbf{x}_0Z^{-1}[\int_{0_{(t)}}^{\infty}dt\int_{0_{(x)}}^{\infty}\int_{0_{(w)}}^{x}e^{-sx}\sum_{i=0}^{\infty}(D_1q_i(t)dH(t+w)A^{[i-1]}(x-w)$$
$$+D_2q_i(t)dA(t+w)A^{[i]}(x-w))$$
$$-D_1\int_{0}^{\infty}q_0(u)dG(u)\cdot\int_{0_{(t)}}^{\infty}dt\int_{0_{(x)}}^{\infty}\int_{0_{(w)}}^{x}e^{-sx}\sum_{i=0}^{\infty}q_i(t)dH(t+w)A^{[i-1]}(x-w)]$$
$$+\sum_{i=1}^{\infty}\sum_{k=1}^{i}\lambda_e\mathbf{x}_k\int_{0_{(t)}}^{t}dt\int_{0_{(x)}}^{\infty}\int_{0_{(w)}}^{x}e^{-sx}q_{i-k}(t)dA(t+w)A^{[i-1]}(x-w)$$
$$=\lambda_e\mathbf{x}_0Z^{-1}$$
$$+\lambda_e\mathbf{x}_0Z^{-1}[\int_{0_{(t)}}^{\infty}dt\int_{0_{(w)}}^{\infty}e^{-sw}\sum_{i=0}^{\infty}\{D_1q_i(t)dH(t+w)A^{*i-1}(s)+D_2q_i(t)dA(t+w)A^{*i}(s)\}$$
$$-D_1\int_{0_{(u)}}^{\infty}q_0(u)dG(u)\cdot\int_{0_{(t)}}^{\infty}dt\int_{0_{(w)}}^{x}e^{-sw}\sum_{i=0}^{\infty}q_i(t)dA(t+w)A^{*i-1}(s)]$$
$$+\sum_{i=1}^{\infty}\sum_{k=1}^{i}\lambda_e\mathbf{x}_k\int_{0_{(t)}}^{\infty}dt\int_{0_{(w)}}^{\infty}e^{-sw}q_{i-k}(t)dA(t+w)A^{*i-1}(s). \quad (3.12)$$

We may assume that arriving non-priority customers in $(0,t)$ receive their services under a preemptive LIFO discipline. Using a similar argument to [12], $\mathbf{V}_{\mathbf{N}}^{*}(s)$ may be written as follows:

$$\mathbf{V}_{\mathbf{N}}^{*}(s)=\lambda_e\mathbf{x}_0Z^{-1}$$
$$+\lambda_e\mathbf{x}_0Z^{-1}\{\int_{0_{(t)}}^{\infty}dt\int_{0_{(w)}}^{\infty}e^{-sw}[D_1dH(t+w)\exp(tL)A^{*-1}(s)+D_2dA(t+w)\exp(tL)]$$
$$-D_1G^*(-M)\int_{0_{(t)}}^{\infty}dt\int_{0_{(w)}}^{\infty}e^{-sw}dA(t+w)\exp(tL)A^{*-1}(s)\}$$
$$+\lambda_e\sum_{k=1}^{\infty}\mathbf{x}_k\int_{0_{(t)}}^{\infty}dt\int_{0_{(w)}}^{\infty}e^{-sw}(dA(t+w))\exp(tL)A^{*k-1}(s). \quad (3.13)$$

Here,

$$L=I_1\otimes T_2+D_2A^*(s) \quad (3.14)$$

and

$$M = I_1 \otimes T_2. \tag{3.15}$$

Now, we obtain

$$\begin{aligned}
&\mathbf{V}^*_{\mathbf{N}}(s) = \lambda_e \mathbf{x}_0 Z^{-1} \\
&+ \lambda_e \mathbf{x}_0 Z^{-1}[D_1 \int_0^\infty e^{-sv} dH(v) \int_0^v \exp\{t(sI+L)\}dt A^{*-1}(s) \\
&+ D_2 \int_0^\infty e^{-sv} dA(v) \int_0^v \exp\{t(sI+L)\}dt \\
&- D_1 G^*(-M) \int_0^\infty e^{-sv} dA(v) \int_0^v \exp\{t(sI+L)\}dt A^{*-1}(s)] \\
&+ \lambda_e \sum_{k=1}^\infty \mathbf{x}_k \int_0^\infty e^{-sv} dA(v) \int_0^v \exp\{t(sI+L)\}dt A^{*k-1}(s).
\end{aligned} \tag{3.16}$$

Thus, we obtain

$$\begin{aligned}
&\mathbf{V}^*_{\mathbf{N}}(s)(sI+L) \\
&= \lambda_e \mathbf{x}_0 Z^{-1}(sI+L) \\
&+ \lambda_e \mathbf{x}_0 Z^{-1}[D_1(H^*(-L) - H^*(s))A^{*-1}(s) + D_2(A^*(-L) - A^*(s)) \\
&- D_1 G^*(-M)(A^*(-L) - A^*(s))A^{*-1}(s)] \\
&+ \lambda_e \sum_{k=1}^\infty \mathbf{x}_k A^*(-L) A^{*k-1}(s) - \lambda_e \sum_{k=0}^\infty \mathbf{x}_k (A^*(s))^k + \lambda_e \mathbf{x}_0.
\end{aligned} \tag{3.17}$$

From (2.20), we have

$$\mathbf{x}_i = \mathbf{x}_0 B_i + \sum_{j=1}^{i+1} \mathbf{x}_j A_{i+1-j},$$

thus,

$$\begin{aligned}
&\sum_{k=0}^\infty \mathbf{x}_k (A^*(s))^k = \mathbf{x}_0 \sum_{i=0}^\infty B_i A^{*i}(s) + \sum_{j=1}^\infty \mathbf{x}_j \sum_{i=0}^\infty A_i A^{*i+j-1}(s) \\
&= \mathbf{x}_0 Z^{-1}[D_1\{G^*(-L) - G^*(-M)\}A^*(-L)A^{*-1}(s) + D_2 A^*(-L)] \\
&+ \sum_{k=1}^\infty \mathbf{x}_k A^*(-L) A^{*k-1}(s).
\end{aligned} \tag{3.18}$$

Substituting this into (3.17), we have

$$\begin{aligned}
&\mathbf{V}^*_{\mathbf{N}}(s)(sI+L) = \lambda_e \mathbf{x}_0 Z^{-1}(sI+L) \\
&+ \lambda_e \mathbf{x}_0 Z^{-1}[D_1\{G^*(-L)A^*(-L) - G^*(s)A^*(s)\}A^{*-1}(s) \\
&+ D_2\{A^*(-L) - A^*(s)\} \\
&- D_1 G^*(-M)\{A^*(-L) - A^*(s)\}A^{*-1}(s) \\
&- D_1\{G^*(-L) - G^*(-M)\}A^*(-L)A^{*-1}(s) - D_2 A^*(-L) - T - D_1 G^*(-M)] \\
&= \lambda_e \mathbf{x}_0 Z^{-1}(sI - T_1 \otimes I_2 - D_1 G^*(s)).
\end{aligned} \tag{3.19}$$

Since

$$\lambda_e \mathbf{x}_0 Z^{-1} = \mathbf{p}_0, \tag{3.20}$$

we have (3.2). ∎

Now, we consider the waiting time distribution of an arriving non-priority customer. We define

$$\mathbf{W_N}(x) = (W_{11}(x), \cdots, W_{1m_2}(x), \cdots, W_{m_1,1}(x), \cdots, W_{m_1,m_2}(x)),$$

where $W_{ij}(x)$ is the joint probability that a non-priority arrival waits at most x before being served and that at arrival, the arrival phases of priority and non-priority customers are i and j, respectively. Here, $\mathbf{W_N}(x)$ is given by

$$\mathbf{W_N}(x) = \lambda_2^{-1}\mathbf{V_N}(x)(I_1 \otimes T_2^0\alpha_2). \tag{3.21}$$

that is,

$$\mathbf{W_N^*}(s) = \lambda_2^{-1}\mathbf{V_N^*}(s)(I_1 \otimes T_2^0\alpha_2).$$

Corollary 3.1

$$\mathbf{g}(z) = \lambda_2^{-1}\int_0^\infty d\mathbf{V_N}(x)\exp\{I_1 \otimes (T_2 + T_2^0\alpha_2 z)x\}(I_1 \otimes T_2^0\alpha_2)C^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z)). \tag{3.22}$$

Proof

From (3.2), we have

$$\begin{aligned}&\int_0^\infty d\mathbf{V_N}(x)\exp\{I_1 \otimes (T_2 + T_2^0\alpha_2 z)x\}\{-I_1 \otimes T_2^0\alpha_2 z + I_1 \otimes T_2^0\alpha_2 \\ &\cdot C^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z))\} \\ &= \lambda_2\mathbf{x}_0\{-T - T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(-I_1 \otimes T_2)\}^{-1}\{-I_1 \otimes T_2^0\alpha_2 z - T - T_1^0\alpha_1 \otimes I_2 \\ &\cdot G^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z))\},\end{aligned} \tag{3.23}$$

thus

$$\begin{aligned}&\int_0^\infty d\mathbf{V_N}(x)\exp\{I_1 \otimes (T_2 + T_2^0\alpha_2 z)x\}(I_1 \otimes T_2^0\alpha_2) \\ &= \lambda_2\mathbf{x}_0\{-T - T_1^0\alpha_1 \otimes I_2 \cdot G_1^*(-I_1 \otimes T_2)\}^{-1}\{-I_1 \otimes T_2^0\alpha_2 z - T - T_1^0\alpha_1 \otimes I_2 \\ &\cdot G^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z))\}\{-zI + C^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z))\}^{-1}.\end{aligned} \tag{3.24}$$

Post-multiplying by $C^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z))$ for both sides of (3.24) and using commutativity of $C^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z))$ and $\{-zI + C^*(I_1 \otimes (-T_2 - T_2^0\alpha_2 z))\}^{-1}$, we obtain (3.22) from (3.21) and (2.18). □

References

[1] Avi-Itzhak, B. and Naor, P., Some queueing problems with the service station subject to breakdown, *Oper. Res.*, **11** (1963) pp. 303–320.

[2] Doshi, B., Generalizations of the stochastic decomposition results for single server queues with vacations, *Comm. Statist. Stochastic Models*, **6** (1990) pp. 307–333.

[3] Gaver, D. P., Junior, A waiting line with interrupted service, including priorities, *J. Roy. Stat. Soc.*, **B 24** (1962) pp. 73–90.

[4] Heffes, H., A class of data traffic processes-Covariance function characterization and related queueing results, *Bell Syst. Tech. J.*, **59** (1980) pp. 897–929.

[5] Heffes, H. and Lucantoni, D. M., A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance, *IEEE J. Selected Areas in Communications*, **SAC-4** (1986) pp. 856–868.

[6] Jaiswal, N. K., *Priority Queues*, Academic Press, New York, 1968.
[7] Keilson, J., Queues subject to service interruption, *Ann. Math. Statist.*, **33** (1962) pp. 1314–1322.
[8] Lucantoni, D. M., Meier-Hellstern, K. S. and Neuts, M. F., A single server queue with server vacations and a class of non-renewal arrival processes, *J. Appl. Prob.*, **22** (1990) pp. 676–705.
[9] Lucantoni, D. M., New results for the single server queue with a batch Markovian arrival process, *Comm. Statist. Stochastic Models*, **7** (1991), pp. 1–46.
[10] Lucantoni, D. M. and Neuts, M. F., Simpler proofs of some properties of the fundamental period of the MAP/G/1 queue, To appear in *J. Appl. Prob.*, April, 1994.
[11] Machihara, F., Completion time of service unit interrupted by *PH*-Markov renewal customers and its application, *Proc. of the 12th Intl. Teletraffic Cong. in Torino*, (1988) 5.4B 5.1–5.8.
[12] Machihara, F., A new approach to the fundamental period of a queue with phase-type Markov renewal arrivals, *Comm. Statist. Stochastic Models*, **6** (1990) pp. 551–560.
[13] Machihara, F., On the queue with *PH*-Markov renewal preemptions, *J. Oper. Res. Soc. Japan*, **36**, 1 (1993) pp. 13–28.
[14] Neuts, M. F., Some explicit formulas for the steady-state behavior of the queue with semi-Markovian service times, *Adv. Appl. Prob.*, **9** (1977) pp. 141–157.
[15] Neuts, M. F., *Matrix Geometric Solutions in Stochastic Models: An Algorithmic Approach*, Baltimore, The Johns Hopkins University Press, 1981.
[16] Neuts, M. F., *Structured Stochastic Matrices of M/G/1 type and Their Applications*, New York and Basel, Marcel Dekker, INC., 1989.
[17] Neuts, M. F., The fundamental period of the queue with Markov modulated arrivals, in *Probability, Statistics and Mathematics: Papers in Honor of Professor Samuel Karlin*, Academic Press, 1989.
[18] Ramaswami, V., The *N/G/1* queue and its detailed analysis, *Adv. Appl. Prob.*, **12** (1980) pp. 222–261.
[19] Ramaswami, V., From the matrix-geometric to the matrix-exponential, *Queueing Systems*, **6** (1990) pp. 229–260.
[20] Sengupta, B., Markov processes whose steady state distribution is matrix-exponential with an application to the *GI/PH/1* queue, *Adv. Appl. Prob.*, **21** (1989) pp. 159–180.
[21] Sengupta, B., The semi-Markovian queue: theory and applications, *Comm. Statist. Stochastic Models*, 6 (1990) pp. 383–413.
[22] Takine, T. and Hasegawa, T., The workload in the MAP/G/1 queue with statedependent services: Its application to a queue with preemptive resume priority, *Comm. Statist. Stochastic Models*, **10** (1994) pp. 183-221.
[23] Welch, P. D., On preemptive resume priority queues, *Ann. Math. Statist.*, 35 (1964) pp. 600–612.
[24] Zhu, Y. and Prabhu, N. U., Markov-modulated PH/G/1 queueing systems, *Queueing System*, (1991) pp. 313–322.

Fumiaki Machihara
NTT Laboratories
3-9-11 Midori-cho, Musashino-shi
Tokyo, 180 Japan
e-mail: fumi@hashi.ntt.jp