Comm. Korean Math. Soc. **14** (1999), No. 4, pp. 805–814

ANALYSIS OF AN MMPP/G/1/K FINITE QUEUE WITH TWO-LEVEL THRESHOLD OVERLOAD CONTROL

GYEMIN LEE AND JONGWOO JEON

ABSTRACT. We consider an MMPP/G/1/K finite queue with two-level threshold overload control. This model has frequently arisen in the design of the integrated communication systems which support a wide range applications having various Quality of Service (QoS) requirements. Through the supplementary variable method, we derive the queue length distribution.

1. Introduction

Significant effort is currently being devoted to the development of integrated communication systems, which can support a wide range applications including voice, video, and data. One of major problem in designing the system is to guarantee various Quality of Service (QoS) requirements of applications.

Once the buffer room is filled up, all incoming messages must be lost. In that time, QoS of all applications present in the system is significantly deteriorated. Therefore overload control, set of all actions reducing the recurrence of such shut-down periods, is an important factor in design of the system.

Overload control is in general performed by selectively discarding cells during each time period where the queue length exceeds a high-level threshold L_2 until it drops to a low-level threshold $L_1(\leq L_2)$. Such a time period is called an *overload period*, and the time interval between two adjacent overload periods is called a *underload period*.

Received July 20, 1999. Revised September 6, 1999.

1991 Mathematics Subject Classification: 60K25, 68M20.

Key words and phrases: overload control, Markov modulated Poisson process, elapsed service time, imbedded Markov chain, supplementary variable method.

Supported by KOSEF, 1996.

Many researchers have studied the overload control in various queueing systems. For the ATM switching with cell discarding scheme, Yegani [5] considered an one-level (i.e., $L_1 = L_2$) overload control in MMPP/G/1 finite queue, where the MMPP is acronym of the Markov modulated Poisson process. He gave a way to derive the loss probability with the embedded Markov chain technique.

In this paper, we focus on a study of the two-level (i.e., $L_1 < L_2$) overload control in an MMPP/G/1 finite queue since the two-level control requires less amount of work to control the system. For the analysis, we employ the supplementary variable method originated by Cox [1].

The rest of this paper is organized as follows. Section 2 consists of the definition of MMPP. In Section 3, we derive the queue length distribution through the supplementary variable method.

2. Markov Modulated Poisson Process

The Markov modulated Poisson process is a doubly stochastic Poisson process, whose arrival rate is a function of the state of a given continuous-time Markov process. We shall denote the infinitesimal generator of the underlying Markov process by Q and the diagonal matrix composed of the arrival rates by Λ.

Define the conditional probabilities

$$P_{i,j}(n,t) = Pr(J(t) = j, N(t) = n | J(0) = i, N(0) = 0)$$

where $N(t)$ and $J(t)$ denote the number of arrivals during $(0,t]$ and the state of the underlying Markov process Q at time t, respectively. In [3], it was showed that matrices $P(n,x) = (P_{i,j}(n,x))_{1 \le i,j \le m}$ have probability generating function

$$\bar{P}(z,t) = \sum_{n=0}^{\infty} z^n P(n,t) = \exp(R(z)t), \ 0 \le z \le 1, \tag{1}$$

with $R(z) = R_0 + zR_1$, where

$$R_0 = Q - \Lambda \text{ and } R_1 = \Lambda.$$

In what follows, we shall assume that the matrix R_0^{-1} exists.

3. Analysis

In this section, we will analyze the MMPP/G/1/K finite queue with two-level overload control through the supplementary variable method. When describing the MMPP, we shall use the same notations as in Section 2. The distribution of a service time will be denoted by $H(x)$, its mean by μ, and its hazard rate function by $r(x)$. During the overload periods, the arrival rate matrix Λ is assumed to be reduced to a given matrix Λ_o.

3.1. Supplementary variable method

Let $X(t)$ denote the number of customers in the system at time t. We shall define the *elapsed service time* $S(t)$ as follows: If $X(t) > 0$, $S(t)$ denotes the amount of service already received by the customer in service at time t. Otherwise, $S(t)$ denotes the amount of time elapsed after the last service completion before t. We also define $O(t)$ to be 'o' if the system stays in overload periods at time t and to be 'u' otherwise. Then, the process $(O(t), J(t), X(t), S(t))$ is a four-dimensional Markov process.

Suppose that

$$\begin{aligned} \pi_u(i,n,x)dx &= \lim_{t\to\infty} Pr(O(t)=u, J(t)=i, X(t)=n, x\le S(t)<x+dx) \\ \pi_o(i,n,x)dx &= \lim_{t\to\infty} Pr(O(t)=o, J(t)=i, X(t)=n, x\le S(t)<x+dx) \end{aligned}$$

exist for all states and define

$$\begin{aligned} \boldsymbol{\pi}_u(n,x) &= (\pi_u(1,n,x),\cdots,\pi_u(m,n,x)), \\ \boldsymbol{\pi}_o(n,x) &= (\pi_o(1,n,x),\cdots,\pi_o(m,n,x)). \end{aligned}$$

Let us recall the definitions of R_0, R_1, $R(z)$, $P(n,x)$, and $\bar{P}(z,x)$ for the uncontrolled arrival rate matrix Λ. In same way, we shall define R_0^o, R_1^o, $R_o(z)$, $P_o(n,x)$, and $\bar{P}_o(z,x)$ for the controlled arrival rate matrix Λ_o. Also, we write

$$P_{u,o}(n,m,x) = \int_0^x P(n-1,y)R_1P_o(m,x-y)dy, \ \ n\ge 1, m\ge 0$$

and denote $P(z,s,x)$ the generating function of $\{P(n,m,x), n\ge 1, m\ge 0\}$.

Conditioning on the state of the process $(O(t), J(t), X(t), S(t))$ time x ago, the Kolmogorov forward equations for the joint distribution $\boldsymbol{\pi}_u(n,x)$

and $\boldsymbol{\pi}_o(n,x)$ can be written down as follows.

$$\begin{aligned}
\boldsymbol{\pi}_u(0,x) &= \boldsymbol{\pi}_u(0,0)P(0,x),\\
\boldsymbol{\pi}_u(n,x) &= \sum_{k=1}^{n} \boldsymbol{\pi}_u(k,0)P(n-k,x)\bar{H}(x),\ 1 \le n \le L_2-1,\\
\boldsymbol{\pi}_o(n,x) &= \sum_{k=1}^{L_2-2} \boldsymbol{\pi}_u(k,0)P_{u,o}(L_2-k,\ n-L_2,x)\bar{H}(x)\cdot 1_{(L_2\le n\le K-1)}\\
&+ \sum_{k=L_1+1}^{n} \boldsymbol{\pi}_o(k,0)P_o(n-k,x)\bar{H}(x),\ L_1+1 \le n \le K-1,\\
\boldsymbol{\pi}_o(K,x) &= \sum_{k=1}^{L_2-2}\sum_{i=K-L_2}^{\infty} \boldsymbol{\pi}_u(k,0)P_{u,o}(L_2-k,i,x)\bar{H}(x)\\
&+ \sum_{k=L_1+1}^{K-1}\sum_{i=K-k}^{\infty} \boldsymbol{\pi}_o(k,0)P_o(i,x)\bar{H}(x).
\end{aligned}$$

where $\bar{H}(x) = 1 - H(x)$ and $1_{(\cdot)}$ is an indicator function. The joint distribution $\boldsymbol{\pi}_u(n,x)$ and $\boldsymbol{\pi}_o(n,x)$ should satisfy the boundary conditions

$$\begin{aligned}
&\boldsymbol{\pi}_u(1,0) = \int_0^\infty \boldsymbol{\pi}_u(2,x)r(x)dx + \int_0^\infty \boldsymbol{\pi}_u(0,x)R_1dx\\
&\boldsymbol{\pi}_u(n,0) = \int_0^\infty \boldsymbol{\pi}_u(n+1,x)r(x)dx,\ 0 \le n \le L_2-2, n \ne 1, L_1\\
&\boldsymbol{\pi}_o(n,0) = \int_0^\infty \boldsymbol{\pi}_o(n+1,x)r(x)dx,\ L_1+1 \le n \le K-1,\\
&\boldsymbol{\pi}_u(L_1,0) = \int_0^\infty \boldsymbol{\pi}_u(L_1+1,x)r(x)dx + \int_0^\infty \boldsymbol{\pi}_o(L_1+1,x)r(x)dx\\
&\boldsymbol{\pi}_u(L_2-1,0) = \mathbf{0}\\
&\boldsymbol{\pi}_o(K,0) = \mathbf{0}
\end{aligned}$$

and the normalization condition

$$\sum_{n=0}^{L_2-1}\int_0^\infty \boldsymbol{\pi}_u(n,x)dx\ \boldsymbol{e} + \sum_{n=L_1+1}^{K}\int_0^\infty \boldsymbol{\pi}_o(n,x)dx\ \boldsymbol{e} = 1.$$

where $\boldsymbol{e} = (1,1,\cdots,1)^t$.

Before finding the coefficients $\boldsymbol{\pi}_u(n,0)$ and $\boldsymbol{\pi}_o(n,0)$ above, we consider the embedded Markov chain $\{O(\tau_n), J(\tau_n), X(\tau_n)\}$, where $\{\tau_n\}$ is the n^{th} epoch of service or idle completion. Then the transition probability matrix of $\{O(\tau_n), J(\tau_n), X(\tau_n)\}$ is

$$Q_E = \begin{pmatrix} F_{u,u} & F_{u,o} \\ F_{o,u} & F_{o,o} \end{pmatrix}$$

where the block $F_{u,u}$ is a square matrix of order $m(L_2-1)$ given by

$$F_{u,u} = \begin{pmatrix} \mathbf{0} & R_1 & \mathbf{0} & \cdots & \mathbf{0} \\ A_0 & A_1 & A_2 & \cdots & A_{L_2-2} \\ \mathbf{0} & A_0 & A_1 & \cdots & A_{L_2-3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & A_0 & A_1 \end{pmatrix}$$

with

$$A_n = \int_0^\infty P(n,x)dH(x).$$

The block $F_{o,o}$ is a square matrix of order $m(K-L_1-1)$ given by

$$\begin{pmatrix} A_1^o & A_2^o & \cdots & A_{K-L_1-2}^o & \sum_{k=K-L_1-1}^\infty A_k^o \\ A_0^o & A_1^o & \cdots & A_{K-L_1-3}^o & \sum_{k=K-L_1-2}^\infty A_k^o \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & A_0^o & \sum_{k=1}^\infty A_1^o \end{pmatrix}$$

with

$$A_n^o = \int_0^\infty P_o(n,x)dH(x).$$

The block $F_{o,u}$ is an $m(K-L_1-1)\times m(L_2-1)$ matrix which is given by

$$\begin{pmatrix} \mathbf{0} & \cdots & \mathbf{0} & A_0^o & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}.$$

The block $F_{u,o}$ is an $m(L_2-1)\times m(K-L_1-1)$ matrix given by

$$\begin{pmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & A_{L_2-1,0} & \cdots & A_{L_2-1,K-L_2-1} & \sum_{k=K-L_2}^{\infty} A_{L_2-1,k} \\ \mathbf{0} & \cdots & \mathbf{0} & A_{L_2-2,0} & \cdots & A_{L_2-2,K-L_2-1} & \sum_{k=K-L_2}^{\infty} A_{L_2-2,k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \cdots & \mathbf{0} & A_{2,0} & \cdots & A_{2,K-L_2-1} & \sum_{k=K-L_2}^{\infty} A_{2,k} \end{pmatrix}$$

with

$$A_{n,m}=\int_0^\infty P_{u,o}(n,m,x)dH(x).$$

THEOREM 1. *The coefficients $\boldsymbol{\pi}_u(n,0)$ and $\boldsymbol{\pi}_o(n,0)$ of joint distribution $\boldsymbol{\pi}_u(n,x)$ and $\boldsymbol{\pi}_o(n,x)$ are given by*

$$\begin{aligned} &(\boldsymbol{\pi}_u(0,0),\cdots,\boldsymbol{\pi}_u(L_2-2,0),\boldsymbol{\pi}_o(L_1+1,0),\cdots,\boldsymbol{\pi}_o(K-1,0)) \\ =\ &\frac{1}{\mu-\boldsymbol{x}_0^u[\mu I+R_0^{-1}]\boldsymbol{e}}(\boldsymbol{x}_0^u,\cdots,\boldsymbol{x}_{L_2-2}^u,\boldsymbol{x}_{L_1+1}^o\cdots,\boldsymbol{x}_{K-1}^o) \end{aligned}$$

where the vector $(\boldsymbol{x}_0^u,\cdots,\boldsymbol{x}_{L_2-2}^u,\boldsymbol{x}_{L_1+1}^o\cdots,\boldsymbol{x}_{K-1}^o)$ is the stationary vector of the matrix Q_E

PROOF. By plugging Kolmogorov forward equations in the boundary conditions, we show that $(\boldsymbol{\pi}_u(0,0),\cdots,\boldsymbol{\pi}_u(L_1-2,0),\boldsymbol{\pi}_o(L_1+1,0),\cdots,\boldsymbol{\pi}_o$ $(K-1,0))$ is a positive invariant vector of the stochastic matrix Q_E, that is,

$$\begin{aligned} &(\boldsymbol{\pi}_u(0,0),\cdots,\boldsymbol{\pi}_u(L_2-2,0),\boldsymbol{\pi}_o(L_1+1,0),\cdots,\boldsymbol{\pi}_o(K-1,0)) \\ =\ &c(\boldsymbol{x}_0^u,\cdots,\boldsymbol{x}_{L_2-2}^u,\boldsymbol{x}_{L_1+1}^o\cdots,\boldsymbol{x}_{K-1}^o) \end{aligned}$$

for some $c>0$. From the normalization condition, we have

$$\boldsymbol{\pi}_u(0,0)[-R_0^{-1}]\boldsymbol{e}+\mu\left[\sum_{n=1}^{L_2-2}\boldsymbol{\pi}_u(n,0)+\sum_{n=L_1+1}^{K-1}\boldsymbol{\pi}_o(n,0)\right]\boldsymbol{e}=1.$$

Therefore, we have

$$c=\frac{1}{\mu-\boldsymbol{x}_0^u[\mu I+R_0^{-1}]\boldsymbol{e}}$$

So the proof is complete. □

The matrices $\mathbf{A}_n$, $\mathbf{A}_n^o$, and $\mathbf{A}_{n,m}$ can be efficiently evaluated by means of the iterative procedure in [2]. See Yegani's results [5] for details. Thus we can compute $\boldsymbol{\pi}_u(n,x)$ and $\boldsymbol{\pi}_o(n,x)$ through deriving the stationary vector of the stochastic matrix Q_E by the standard methods (e.g., Gauss-Shield algorithm, etc.).

3.2. Queue length distribution

In this subsection, we shall give a way to obtain the queue length distributions $\boldsymbol{\pi}_u(n)$ and $\boldsymbol{\pi}_o(n)$ using the coefficients $\boldsymbol{\pi}_u(n,0)$ and $\boldsymbol{\pi}_o(n,0)$ derived in the previous subsection.

Let us first define

$$\begin{aligned} M_n &= \int_0^\infty P(n,x)\bar{H}(x)dx, \ n \geq 0 \\ M_n^o &= \int_0^\infty P_o(n,x)\bar{H}(x)dx, \ n \geq 0 \\ M_{n,m} &= \int_0^\infty P_o(n,m,x)\bar{H}(x)dx, \ n \geq 1, m \geq 0 \end{aligned}$$

and denote by $M(z)$, $M_o(z)$, and $M(z,s)$ the generating functions of $\{M_n, n \geq 0\}$, $\{M_n^o, n \geq 0\}$, and $\{M_{n,m}, n \geq 1, m \geq 0\}$. Then Kolmogorov forward equations yield

$$\boldsymbol{\pi}_u(0) = \boldsymbol{\pi}_u(0,0)[-R_0^{-1}] \tag{2}$$

$$\boldsymbol{\pi}_u(n) = \sum_{k=1}^{n} \boldsymbol{\pi}_u(k,0)M_{n-k}, \ 1 \leq n \leq L_2 - 1, \tag{3}$$

$$\begin{aligned} \boldsymbol{\pi}_o(n) &= \sum_{k=1}^{L_2-2} \boldsymbol{\pi}_u(k,0)M_{L_2-k,n-L_2} \cdot 1_{(L_2 \leq n \leq K-1)} \\ &+ \sum_{k=L_1+1}^{n} \boldsymbol{\pi}_o(k,0)M_{n-k}^o, \ L_1 + 1 \leq n \leq K-1. \end{aligned} \tag{4}$$

From the fact that $\sum_{n=0}^{L_2-1} \boldsymbol{\pi}_u(n) + \sum_{n=L_1+1}^{K} \boldsymbol{\pi}_o(n)$ is the stationary vector of the underlying Markov process Q, we get

$$\boldsymbol{\pi}_o(K) = \boldsymbol{\theta} - \sum_{n=0}^{L_2-1} \boldsymbol{\pi}_u(n) - \sum_{n=L_1+1}^{K-1} \boldsymbol{\pi}_o(n) \tag{5}$$

where $\boldsymbol{\theta}$ is the stationary vector of the matrix Q. Therefore, we can obtain the queue length distribution if the matrices M_n, M_n^o, and $M_{n,m}^{u,o}$ are given.

From the arguments in [4], the matrices M_n and M_n^o can be obtained in the following way:

$$\begin{aligned} A_n &= M_n R_0 + M_{n-1} R_1 + 1_{(n=0)} \cdot I, \\ A_n^o &= M_n^o R_0^o + M_{n-1}^o R_1^o + 1_{(n=0)} \cdot I. \end{aligned}$$

Similarly, the following lemma enables us to obtain $M_{n,m}$ from $A_{n,m}$ recursively.

LEMMA 1. *The matrices $M_{n,m}$ satisfy*

$$\begin{aligned} A_{n,0} &= M_{n-1} R_1 + M_{n,0} R_0^o, \ n \geq 1 \\ A_{n,m} &= M_{n,m} R_0^o + M_{n,m-1} R_1^o, \ n \geq 1, m \geq 1. \end{aligned}$$

PROOF. Differentiating the generating function $P_{u,o}(z,s,x)$ of $\{P_{u,o}(n,m,x), n \geq 1, m \geq 0\}$, we have

$$\begin{aligned} & \frac{d}{dx}\left(\int_0^x zP(z,y)R_1 P_o(s,x-y)dy\right) \\ = \ & zP(z,x)R_1 + \int_0^x zP(z,y)R_1 \frac{d}{dx} P_o(s,x-y)dy \\ = \ & zP(z,x)R_1 + P(z,s,x)R_o(s). \end{aligned}$$

The first equality above is obtained by Leibniz's theorem and the fact that $P_o(z,0) = I$. The second equality is derived from the equation (1). If we integrate the generating function $A(z,s)$ of $\{A_{n,m}, n \geq 1, m \geq 0\}$ by part, we have

$$\begin{aligned} A(z,s) &= \int_0^\infty P(z,s,x)dH(x) \\ &= \int_0^\infty \frac{d}{dx} P(z,s,x)\bar{H}(x)dx \\ &= zM(z)R_1 + M(z,s)R^o(s). \end{aligned}$$

With the above equation, we directly derive this lemma. So the proof is complete. □

From the above relation of $(M_n, M_n^o, M_{n,m})$ and $(A_n, A_n^o, A_{n,m})$ and the equations (2)-(4), we can also derive the queue length distribution $\boldsymbol{\pi}_u(n)$

and $\pi_o(n)$ in the following way.

$$\begin{aligned}
\pi_u(0) &= \pi_u(0,0)[-R_0]^{-1} \\
\pi_u(1) &= [\pi_u(1,0) - \pi_u(0,0)][-R_0]^{-1} \\
\pi_u(n) &= [\pi_u(n,0) - \pi_u(n-1,0)][-R_0]^{-1} + \pi_u(n-1)R_1[-R_0]^{-1}, \\
&\quad 2 \le n \le L_2 - 1, n \ne L_1 + 1 \\
\pi_o(n) &= [\pi_o(n,0) - [\pi_o(n-1,0)][-R_0^o]^{-1} + \pi_o(n-1)R_1^o[-R_0^o]^{-1} \\
&\quad L_1 + 1 < n \le K - 1, n \ne L_2
\end{aligned}$$

and

$$\begin{aligned}
\pi_u(n) &= [\pi_u(n,0) - \pi_u(n-1,0)][-R_0]^{-1} + \pi_u(n-1)R_1[-R_0]^{-1} \\
&+ \pi_o(n,0)A_0^o[-R_0^{-1}],\ n = L_1 + 1 \\
\pi_o(n) &= \pi_o(n,0)[I - A_0^o][-R_0^o]^{-1},\ n = L_1 + 1 \\
\pi_o(n) &= [\pi_o(n,0) - \pi_o(n-1,0)][-R_0^o]^{-1} + \pi_o(n-1)R_1^o[-R_0^o]^{-1} \\
&+ \pi_u(n-1)R_1[-R_0^o]^{-1},\ n = L_2.
\end{aligned}$$

Thus we can obtain $\pi_u(n)$ and $\pi_o(n)$ without derivation of M_n, M_n^o, and $M_{n,m}$.

References

[1] D. R. Cox, *The analysis of non-Markovian stochastic process by inclusion of supplementary variables*, Proc. Camb. Phil. Soc. **51** (1955), 433-441.

[2] D. M. Lucantoni and V. Ramaswam, *Efficient algorithms for solving the non linear matrix equation arising in phase type queues*, Commun. Statist.-Stochastic Model **1** (1985), 29-52.

[3] M. F. Neuts, *A versatile Markovian point process*, J. Appl. Prob. **16** (1979), 764-779.

[4] V. Ramaswami, *The N/G/1 queue and its detailed analysis*, Adv. Appl. Prob. **12** (1980), 222-261.

[5] P. Yegani, *Performance Models for ATM Switching of Mixed Continuous -Bit-Rate and Bursty Traffic with Threshold-Based Discarding*, IEEE ICC'93, 354.3.1-354.3.7.

Gyemin Lee
Department of statistics
Gyeongsang National University
Chinju 660-701, Korea

Jongwoo Jeon
Department of statistics
Seoul National University
Seoul 151-742, Korea