Mathematical Problems in Engineering
Volume 5, pp. 329–348
Reprints available directly from the publisher
Photocopying permitted by license only

A Queueing Model for Error Control of Partial Buffer Sharing in ATM

BOO YONG AHN and HO WOO LEE*

School of Systems Management Engineering, Sung Kyun Kwan University, Su Won, Republic of Korea 440-746

(Received 13 October 1998)

We model the error control of the partial buffer sharing of ATM by a queueing system $M_1, M_2/G/1/K+1$ with threshold and instantaneous Bernoulli feedback. We first derive the system equations and develop a recursive method to compute the loss probabilities at an arbitrary time epoch. We then build an approximation scheme to compute the mean waiting time of each class of cells. An algorithm is developed for finding the optimal threshold and queue capacity for a given quality of service.

Keywords: ATM; Partial buffer sharing; Error control; Loss probability; Waiting time

1. INTRODUCTION

Asynchronous Transfer Mode (ATM) is considered as the basis of the packet-switching technologies of the next generation [3]. ATM is a packet transfer mode using fixed size information cells. ATM cells have a cell loss priority bit in the header such that high- and low-priority classes can be distinguished. Numerous researchers have focused on developing priority control scheme for efficient buffer utilization (Bae and Suda [2]). Those works show that it is primarily important to design a buffer control scheme such that the control policy satisfies the quality of service (QoS) for different priority classes. Usually the delay-sensitive data are given low priority and the loss-sensitive data are classified as high-priority class.

* Corresponding author. E-mail: hwlee@yurim.skku.ac.kr.

Loss priority control is concerned with reducing the cell loss probability of the loss-sensitive data. Studies on loss priority control can be found in Garcia and Casals [5], Kröner [10], Hebuterne and Gravey [8], Petr and Frost [13], Suri *et al.* [14], Chang and Tan [4], and Akyildiz and Cheng [1]. Kröner [10] proposed two mechanisms for controlling the loss priority: "push-out mechanism" and "partial buffer sharing mechanism".

In the push-out mechanism, if the buffer is full and a high-priority cell (type-1 customer) arrives, a cell with low priority (type-2 customer) is pushed out and lost. In the partial buffer sharing mechanism, low-priority cells can only access the buffer if the buffer occupancy is less than a given integral value T. Kröner [10] showed that the system performance can be improved by using priorities and that the partial buffer sharing mechanism is a good compromise between performance and implementation.

In order to achieve a reliable data transmission in ATM networks, each terminal should have some error recovery scheme. There are two methods for the error recovery: "retransmission method" and "forward error correction method" [12]. In retransmission method, if the receiver receives a correct cell, it returns an ACKnowledgment (ACK) to the sender. Otherwise, it discards the incorrect cell and returns a Negative ACKnowledgment (NACK). The sender upon receipt of NACK retransmits the cell. In foward error correction method, the receivers recovers the lost cells by using a coding technique [18].

In this paper, we build an $M_1, M_2/G/1/K+1$ queueing system with threshold and instantaneous Bernoulli feedback for the optimal control of the partial buffer sharing in ATM. As an error recovery scheme, we assume the retransmission method. We use a recursive method [6,7,11,20] to compute the state probabilities and performance measures.

2. THE MODEL

2.1. The System

We consider the queueing system with following specifications (Fig. 1):

(1) High-priority cells (type-1 cells) and low-priority cells (type-2 cells) arrive singly according to independent Poisson processes with rates λ_1 and λ_2.

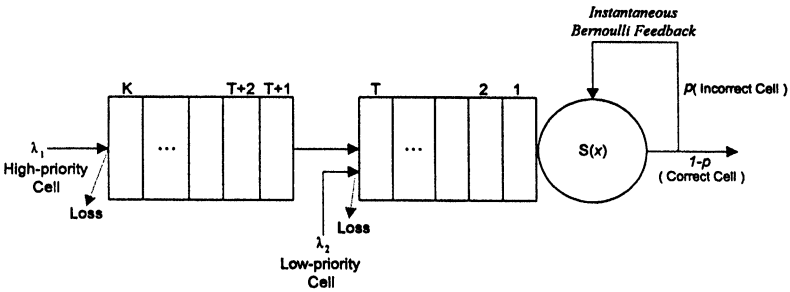

FIGURE 1 The system.

(2) The system has a single processor and K waiting spaces (therefore there can be $K+1$ cells in the system including the one in processing).
(3) Processing times are assumed to be identically and generally distributed.
(4) Queue disciplines are as follows:
   (i) If the queue length (i.e., the number of cells in the queue excluding the one under processing) is less than the threshold T, all arriving cells enter the system.
   (ii) If the queue length is greater than or equal T, only type-1 cells are allowed to enter the system.
   (iii) If the buffer is full, all arriving cells are lost.
(5) After being served, a cell leaves the system with probability $1-p$ or rejoins the server instantaneously with probability p.

2.2. System Equations

Let us define the following notations and probabilities:

λ_1	arrival rate of high-priority (type-1) cells,
λ_2	arrival rate of low-priority (type-2) cells,
λ	total arrival rate of cells ($=\lambda_1+\lambda_2$),
p	feedback probability of a cell,
T	threshold,
K	queue capacity,
$S(x)$	processing time distribution function,
$s(x)$	processing time probability density function,

$S^*(\theta)$	Laplace transform of $s(x)$,
$E(S)$	mean processing time,
$Var(S)$	variance of S,
$P_{\text{Loss }1}$	loss probability of type-1 cell,
$P_{\text{Loss }2}$	loss probability of type-2 cell,
a	offered load $(=(\lambda_1+\lambda_2)E(S))$,
a'	carried load $(=[\lambda_1(1-P_{\text{Loss }1})+\lambda_2(1-P_{\text{Loss }2})]\,E(S))$,
$R(t)$	remaining processing time of the cell under processing at time t,
$N(t)$	total number of cells in the system (including the one in processing) at time t,

$$P_0(t) = Pr[N(t)=0],$$
$$P_n(x,t)\Delta x = Pr[N(t)=n,\ x<R(t)\le x+\Delta x],\ (n=1,\dots,K+1),$$
$$P_n(t) = Pr[N(t)=n] = \textstyle\int_0^\infty P_n(x,t)\,\mathrm{d}x,\ (n=1,\dots,K+1).$$

It is not difficult to derive the following steady-state system equations:

$$0 = -\lambda P_0 + (1-p)P_1(0), \tag{2.1}$$

$$-\frac{\mathrm{d}}{\mathrm{d}x}P_1(x) = -\lambda P_1(x) + ps(x)P_1(0) + (1-p)s(x)P_2(0) + \lambda P_0 s(x), \tag{2.2}$$

$$-\frac{\mathrm{d}}{\mathrm{d}x}P_n(x) = -\lambda P_n(x) + ps(x)P_n(0) + (1-p)s(x)P_{n+1}(0) + \lambda P_{n-1}(x), \quad (2\le n\le T), \tag{2.3}$$

$$-\frac{\mathrm{d}}{\mathrm{d}x}P_{T+1}(x) = -\lambda_1 P_{T+1}(x) + ps(x)P_{T+1}(0) + (1-p)s(x)P_{T+2}(0) + \lambda P_T(x), \tag{2.4}$$

$$-\frac{\mathrm{d}}{\mathrm{d}x}P_n(x) = -\lambda_1 P_n(x) + ps(x)P_n(0) + (1-p)s(x)P_{n+1}(0) + \lambda_1 P_{n-1}(x), \quad (T+2\le n\le K), \tag{2.5}$$

$$-\frac{\mathrm{d}}{\mathrm{d}x}P_{K+1}(x) = ps(x)P_{K+1}(0) + \lambda_1 P_K(x). \tag{2.6}$$

Let us define the following Laplace transform:

$$P_n^*(\theta) = \int_0^\infty \mathrm{e}^{-\theta x} P_n(x), \quad (1 \leq n \leq K+1).$$

Taking the Laplace transforms of both sides of (2.2)–(2.6), we obtain the following transform equations:

$$(\lambda - \theta)P_1^*(\theta) = S^*(\theta)[\lambda P_0 + pP_1(0) + (1-p)P_2(0)] - P_1(0), \quad (2.7)$$

$$(\lambda - \theta)P_n^*(\theta) = \lambda P_{n-1}^*(\theta) + S^*(\theta)[pP_n(0) + (1-p)P_{n+1}(0)] - P_n(0), \quad (n = 2, \ldots, T), \quad (2.8)$$

$$(\lambda_1 - \theta)P_{T+1}^*(\theta) = \lambda P_T^*(\theta) + S^*(\theta)[pP_{T+1}(0) + (1-p)P_{T+2}(0)] - P_{T+1}(0), \quad (2.9)$$

$$(\lambda_1 - \theta)P_n^*(\theta) = \lambda_1 P_{n-1}^*(\theta) + S^*(\theta)[pP_n(0) + (1-p)P_{n+1}(0)] - P_n(0), \quad (n = T+2, \ldots, K), \quad (2.10)$$

$$-\theta P_{K+1}^*(\theta) = \lambda_1 P_K^*(\theta) + pS^*(\theta)P_{K+1}(0) - P_{K+1}(0). \quad (2.11)$$

2.3. Probability Computation

Let the steady-state probability be

$$P_n = P_n^*(0) = P_n^*(\theta)|_{\theta=0} = \int_0^\infty P_n(x)\,\mathrm{d}x, \quad (n = 1, \ldots, K+1).$$

From (2.1), we have

$$P_1(0) = \lambda P_0/(1-p).$$

Using (2.1) in (2.7) and adding (2.7)–(2.11), we get

$$\sum_{n=1}^{K+1} P_n^*(\theta) = \left[\frac{1 - S^*(\theta)}{\theta}\right] \sum_{n=1}^{K+1} P_n(0). \quad (2.12)$$

From (2.12) and using L'hospital's rule, we get

$$\sum_{n=1}^{K+1} P_n^*(0) = \left[\frac{-S^{*(1)}(\theta)}{1}\sum_{n=1}^{K+1} P_n(0)\right]\Bigg|_{\theta=0} = E(S)\sum_{n=1}^{K+1} P_n(0). \quad (2.13)$$

Using (2.1) in (2.7) and letting $\theta = \lambda$ and $\theta = 0$ respectively in (2.7), we get

$$P_2(0) = \frac{\lambda}{(1-p)^2}\left[\frac{1-S^*(\lambda)}{S^*(\lambda)}\right]P_0, \quad (2.14)$$

$$P_1^*(0) = \frac{(1-p)}{\lambda}P_2(0). \quad (2.15)$$

Note that $S^*(\lambda)$ is the probability that no cells arrive during a processing time. Now letting $\theta = \lambda$ in (2.8), we obtain

$$P_{n+1}(0) = \frac{[1-pS^*(\lambda)]P_n(0) - \lambda P_{n-1}^*(\lambda)}{(1-p)S^*(\lambda)}, \quad (n = 2, \ldots, T). \quad (2.16)$$

Let us define jth derivative of $P_n^*(\theta)$ as

$$\frac{\mathrm{d}^j P_n^*(\theta)}{\mathrm{d}\theta^j} = P_n^{*(j)}(\theta), \quad (n = 1, \ldots, K+1).$$

From (2.7) and (2.8), we obtain

$$P_1^{*(j)}(\lambda) = -\frac{1}{j+1}S^{*(j+1)}(\lambda)[P_1(0) + (1-p)P_2(0)],$$
$$(j = 0, \ldots, T-2), \quad (2.17)$$

$$P_n^{*(j)}(\lambda)$$
$$= -\frac{1}{j+1}\left\{\lambda P_{n-1}^{*(j+1)}(\lambda) + S^{*(j+1)}(\lambda)[pP_n(0) + (1-p)P_{n+1}(0)]\right\},$$
$$(n = 2, \ldots, T,\ j = 0, \ldots, T-n-1), \quad (2.18)$$

where $P_n^{*(0)}(\lambda) = P_n^*(\lambda)$. Now $P_n(0)$ $(3 \le n \le T+1)$ can be expressed in terms of P_0 recursively from (2.17) and (2.18).

Letting $\theta = \lambda_1$ in (2.9), we get

$$P_{T+2}(0) = \frac{[1 - pS^*(\lambda_1)]P_{T+1}(0) - \lambda P_T^*(\lambda_1)}{(1-p)S^*(\lambda_1)}. \tag{2.19}$$

From (2.7) and (2.8), we get

$$P_1^*(\lambda_1) = \frac{S^*(\lambda_1)[P_1(0) + (1-p)P_2(0)] - P_1(0)}{\lambda_2}, \tag{2.20}$$

$$P_n^*(\lambda_1) = \frac{\lambda P_{n-1}^*(\lambda_1) + S^*(\lambda_1)[pP_n(0) + (1-p)P_{n+1}(0)] - P_n(0)}{\lambda_2},$$
$$(n = 2, \dots, T). \tag{2.21}$$

Now, $P_{T+2}(0)$ can be expressed in terms of P_0 recursively from (2.20) and (2.21).

Letting $\theta = \lambda_1$ in (2.10), we get

$$P_{n+1}(0) = \frac{[1 - pS^*(\lambda_1)]P_n(0) - \lambda_1 P_{n-1}^*(\lambda_1)}{(1-p)S^*(\lambda_1)}, \quad (n = T+2, \dots, K). \tag{2.22}$$

From (2.7)–(2.10), we get

$$P_1^{*(j)}(\lambda_1) = \frac{S^{*(j)}(\lambda_1)[P_1(0) + (1-p)P_2(0)] + jP_1^{*(j-1)}(\lambda_1)}{\lambda_2},$$
$$(j = 1, \dots, K-T), \tag{2.23}$$

$$P_n^{*(j)}(\lambda_1) = \frac{\lambda P_{n-1}^{(j)}(\lambda_1) + S^{*(j)}(\lambda_1)[pP_n(0) + (1-p)P_{n+1}(0)] + jP_n^{*(j-1)}(\lambda_1)}{\lambda_2}$$
$$(n = 2, \dots, T,\ j = 1, \dots, K-T), \tag{2.24}$$

$$P_{T+1}^{*(j)}(\lambda_1) = -\frac{1}{j+1}\left\{\lambda P_T^{*(j+1)}(\lambda_1) + S^{*(j+1)}(\lambda_1)[pP_{T+1}(0) + (1-p)P_{T+2}(0)]\right\},$$
$$(j = 0, \dots, K-T-2), \tag{2.25}$$

$$P_n^{*(j)}(\lambda_1) = -\frac{1}{j+1}\left\{\lambda_1 P_{n-1}^{*(j+1)}(\lambda_1) + S^{*(j+1)}(\lambda_1)[pP_n(0) + (1-p)P_{n+1}(0)]\right\},$$
$$(n = T+2,\ldots,K-1,\ j=0,\ldots,K-n-1). \tag{2.26}$$

Hence $P_{T+1}^{*(j)}(\lambda_1)$ $(j=0,\ldots,K-T-2)$ can be obtained recursively from (2.23) and (2.24). Also $P_n(0)$ $(T+3 \le n \le K+1)$ can be obtained in terms of P_0 recursively from (2.25) and (2.26).

Letting $\theta = 0$ in (2.8)–(2.10), we get

$$P_n^*(0) = \frac{\lambda P_{n-1}^*(0) + (1-p)[P_{n+1}(0) - P_n(0)]}{\lambda}, \quad (n = 2,\ldots,T), \tag{2.27}$$

$$P_{T+1}^*(0) = \frac{\lambda P_T^*(0) + (1-p)[P_{T+2}(0) - P_{T+1}(0)]}{\lambda_1}, \tag{2.28}$$

$$P_n^*(0) = \frac{\lambda_1 P_{n-1}^*(0) + (1-p)[P_{n+1}(0) - P_n(0)]}{\lambda_1}, \quad (n = T+2,\ldots,K). \tag{2.29}$$

The only unknown quantity $P_{K+1}^*(0)$ can be obtained from (2.13),

$$P_{K+1}^*(0) = E(S)\sum_{n=1}^{K+1} P_n(0) - \sum_{n=1}^{K} P_n^*(0). \tag{2.30}$$

P_0 can be obtained from the normalization condition and is given by

$$P_0 = \left[1 + \sum_{n=1}^{K+1} P_n^*(0)\right]^{-1} \tag{2.31}$$

The above procedure to compute $\{P_n,\ n=0,\ldots,K+1\}$ can be summarized as follows:

Step 1: Let $P_0 = 1$ (P_0 will be calibrated after the normalization).
Compute $P_1(0)$ from (2.1).
Compute $P_2(0)$ from (2.14).
Compute $P_1^*(0)$ from (2.15).

Step 2: Compute $P_1^{*(j)}(\lambda)$, $(j=0,\ldots,T-2)$ from (2.17).
Compute $P_n^{*(j)}(\lambda)$, $(n=2,\ldots,T, j=0,\ldots,T-n-1)$ from (2.18).

Compute $P_{n+1}(0)$, $(n = 2, \ldots, T)$ from (2.16).
Compute $P_1^*(\lambda_1)$ from (2.20).
Compute $P_n^*(\lambda_1)$, $(n = 2, \ldots, T)$ from (2.21).
Compute $P_{T+2}(0)$ from (2.19).
Compute $P_1^{*(j)}(\lambda_1)$, $(j = 1, \ldots, K - T)$ from (2.23).
Compute $P_n^{*(j)}(\lambda_1)$, $(n = 2, \ldots, T, j = 1, \ldots, K - T)$ from (2.24).
Compute $P_{T+1}^{*(j)}(\lambda_1)$, $(j = 0, \ldots, K - T - 2)$ from (2.25).
Compute $P_n^{*(j)}(\lambda_1)$, $(n = T + 2, \ldots, K - 1, j = 0, \ldots, K - n - 1)$ from (2.26).
Compute $P_{n+1}(0)$, $(n = T + 2, \ldots, K)$ from (2.22).

Step 3: Compute $P_n^*(0)$, $(n = 2, \ldots, T)$ from (2.27).
Compute $P_{T+1}^*(0)$ from (2.28).
Compute $P_n^*(0)$, $(n = T + 2, \ldots, K)$ from (2.29).

Step 4: Compute $P_{K+1}^*(0)$ from (2.30).

Step 5: Compute $SUM = 1 + \sum_{n=1}^{K+1} P_n^*(0)$.
Compute $P_0 = 1/SUM$.
Compute $P_n = P_n^*(0)/SUM$, $(n = 1, \ldots, K + 1)$.

2.4. Performance Measures

Let us define the following notations:

L	mean number of cells (combining both types),
L_q	mean queue size (combining both types),
W	mean system sojourn time (irrespective of cell types),
W_q	mean queue waiting time (irrespective of cell types),
λ_1'	effective arrival rate of high-priority cell ($= \lambda_1(1 - P_{\text{Loss 1}})$),
λ_2'	effective arrival rate of low-priority cell ($= \lambda_2(1 - P_{\text{Loss 2}})$),
λ'	effective total arrival rate ($= \lambda_1' + \lambda_2'$).

Then we have

$$L = \sum_{n=0}^{K+1} n \cdot P_n \tag{2.32}$$

$$L_q = \sum_{n=1}^{K+1} (n - 1) \cdot P_n \tag{2.33}$$

From Little's law, we get

$$W = L/\lambda', \qquad W_q = L_q/\lambda'. \tag{2.34}$$

Let $P_{\text{Loss}\,1}$ and $P_{\text{Loss}\,2}$ be the loss probabilities of each types of cells respectively. Then from PASTA (Poisson Arrivals See Time Average, Wolff [19]), we get

$$P_{\text{Loss}\,1} = P_{K+1}, \tag{2.35}$$

$$P_{\text{Loss}\,2} = \sum_{n=T+1}^{K+1} P_n. \tag{2.36}$$

2.5. Numerical Example

We consider an example with $\lambda_1 = 1.0$, $\lambda_2 = 2.0$, $p = 0.01$, $T = 3$, $K = 5$. We assume deterministic processing time with $E(S) = 0.3$. Then we have $\lambda = \lambda_1 + \lambda_2 = 3.0$, $S^*(\theta) = \mathrm{e}^{-0.3\theta}$.

Step 1: Compute $P_n(0)$, $(n = 1, 2)$ and $P_1^*(0)$.

$P_0(0) = 1.0,$
$P_1(0) = \lambda P_0/(1-p) = 3.030303\,03,$
$P_2(0) = \lambda P_0[1 - S^*(\lambda)]/[(1-p)^2 S^*(\lambda)] = 4.467716\,899,$
$P_1^*(0) = (1-p)P_2(0)/\lambda = 1.474346577.$

Step 2: Compute $P_n(0)$, $(n = 3, 4, 5, 6)$.

$P_1^*(\lambda) = -S^{*(1)}(\lambda)[P_1(0) + (1-p)P_2(0)] = 0.909090\,909,$
$P_1^{*(1)}(\lambda) = -S^{*(2)}(\lambda)[P_1(0) + (1-p)P_2(0)]/2 = 0.136363\,636,$
$P_3(0) = \{[1 - pS^*(\lambda)]P_2(0) - \lambda P_1^*(\lambda)\}/[(1-p)S^*(\lambda)]$
$\quad = 4.278913\,872,$
$P_2^*(\lambda) = -\{\lambda P_1^{*(1)}(\lambda) + S^{*(1)}(\lambda)[pP_2(0)$
$\quad +(1-p)P_3(0)]\} = 0.931224\,16,$
$P_4(0) = \{[1 - pS^*(\lambda)]P_3(0) - \lambda P_2^*(\lambda)\}/[(1-p)S^*(\lambda)]$
$\quad = 3.646783\,036,$
$P_1^*(\lambda_1) = \{S^*(\lambda_1)[P_1(0) + (1-p)P_2(0)] - P_1(0)\}/\lambda_2$
$\quad = 1.245634\,547,$
$P_2^*(\lambda_1) = \{\lambda P_1^*(\lambda_1) + S^*(\lambda_1)[pP_2(0)$
$\quad +(1-p)P_3(0)] - P_2(0)\}/\lambda_2 = 1.220241\,396,$
$P_3^*(\lambda_1) = \{\lambda P_2^*(\lambda_1) + S^*(\lambda_1)[pP_3(0)$
$\quad +(1-p)P_4(0)] - P_3(0)\}/\lambda_2 = 1.044048\,289,$

$$P_5(0) = \{[1 - pS^*(\lambda_1)]P_4(0) - \lambda P_3^*(\lambda_1)\}/[(1-p)S^*(\lambda_1)]$$
$$= 0.664869\,732,$$
$$P_1^{*(1)}(\lambda_1) = \{S^{*(1)}(\lambda_1)[P_1(0) + (1-p)P_2(0)]$$
$$+P_1^*(\lambda_1)\}/\lambda_2 = -0.20541\,8545,$$
$$P_1^{*(2)}(\lambda_1) = \{-S^{*(2)}(\lambda_1)[P_1(0)+(1-p)P_2(0)] + P_1^{*(1)}(\lambda_1)\}/\lambda_2$$
$$= 0.043052\,201,$$
$$P_2^{*(1)}(\lambda_1) = \{\lambda P_1^{*(1)}(\lambda_1) + S^{*(1)}(\lambda_1)[pP_2(0) + (1-p)P_3(0)]$$
$$+P_2^*(\lambda_1)\}/\lambda_2 = -0.17370\,1527,$$
$$P_3^{*(1)}(\lambda_1) = \{\lambda P_2^{*(1)}(\lambda_1) + S^{*(1)}(\lambda_1)[pP_3(0) + (1-p)P_4(0)]$$
$$+P_3^*(\lambda_1)\}/\lambda_2 = -0.14447\,1085,$$
$$P_4^*(\lambda_1) = -\{\lambda P_3^{*(1)}(\lambda_1) + S^{*(1)}(\lambda_1)[pP_4(0) + (1-p)P_5(0)]\}$$
$$= 0.587804\,706,$$
$$P_6(0) = \{[1 - pS^*(\lambda_1)]P_5(0) - \lambda_1 P_4^*(\lambda_1)\}/[(1-p)S^*(\lambda_1)]$$
$$= 0.098361\,825.$$

Step 3: Compute $P_n^*(0)$, $(n=2,3,4,5)$.

$$P_2^*(0) = \{\lambda P_1^*(0) + (1-p)[P_3(0) - P_2(0)]\}/\lambda$$
$$= 1.412041\,578,$$
$$P_3^*(0) = \{\lambda P_2^*(0) + (1-p)[P_4(0) - P_3(0)]\}/\lambda$$
$$= 1.203438\,402,$$
$$P_4^*(0) = \{\lambda P_3^*(0) + (1-p)[P_5(0) - P_4(0)]\}/\lambda_1$$
$$= 0.658221\,035,$$
$$P_5^*(0) = \{\lambda_1 P_4^*(0) + (1-p)[P_6(0) - P_5(0)]\}/\lambda_1$$
$$= 0.097378\,207.$$

Step 4: Compute $P_6^*(0)$.

$$P_6^*(0) = E(S)\textstyle\sum_{n=1}^{6} P_n(0) - \sum_{n=1}^{5} P_n^*(0) = 0.010658719.$$

Step 5: Compute P_n, $(n=0,1,\ldots,6)$.

$$SUM = 1 + \textstyle\sum_{n=1}^{6} P_n^*(0) = 5.856084\,518,$$
$$P_0 = 1/SUM = 0.170762\,562,$$
$$P_1 = P_1^*(0)/SUM = 0.251763\,199,$$
$$P_2 = P_2^*(0)/SUM = 0.241123\,837,$$
$$P_3 = P_3^*(0)/SUM = 0.205502\,225,$$
$$P_4 = P_4^*(0)/SUM = 0.112399\,511,$$
$$P_5 = P_5^*(0)/SUM = 0.016628\,552,$$
$$P_6 = P_6^*(0)/SUM = 0.001820\,110.$$

Then we have $P_{\text{Loss}\,1} = P_6 = 0.001820\,110$, $P_{\text{Loss}\,2} = \sum_{n=4}^{6} = 0.130848\,172$.

2.6. Approximation of the Mean Sojourn Time of Each Class

In this section we derive an approximate mean waiting time and system sojourn time of each class of cells following the method described in Lee and Ahn [11]. Let us define the following notations:

P_n probability that an arriving cell sees n cells in the system,
$T_{R,n}$ remaining total processing time of the cell under processing when the arriving test cell finds n cells in the system,
$W_q^{(1)}$ approximate mean queue waiting time of a high-priority (type-1) cell,
$W_q^{(2)}$ approximate mean queue waiting time of a low-priority (type-2) cell,
$W^{(1)}$ approximate mean system sojourn time (=queue waiting time + processing time) of a high-priority cell,
$W^{(2)}$ approximate mean system sojourn time (=queue waiting time + processing time) of a low-priority cell,
Q_n^1 probability that the position of the high-priority test cell is n given that it enters the system, $(n=1,\dots,K+1)$,
Q_n^2 probability that the position of the low-priority test cell is n given that it enters the system, $(n=1,\dots,T+1)$.

From PASTA, we have

$$Q_n^1 = \frac{P_{n-1}}{\sum_{m=0}^{K} P_m} = \frac{P_{n-1}}{(1-P_{K+1})} = \frac{P_{n-1}}{(1-P_{\text{Loss}\,1})}, \quad (n=1,\dots,K+1) \tag{2.37}$$

Similarly, we have

$$Q_n^2 = \frac{P_{n-1}}{(1-P_{\text{Loss}\,2})}, \quad (n=1,\dots,T+1) \tag{2.38}$$

We first note that the number of cells that is found by an arriving cell and the remaining total processing time of the in-service cell are not independent. Takagi [16] shows how to obtain the joint transform of the two quantities in case of simple finite capacity $M/G/1$ queue with one class of customers. It is expected that the joint transform for two class cases is more difficult to obtain. With this difficulty in mind, we proceed to obtain the approximate mean waiting time as follows:

Let the total processing time of a cell be S_T. Then we have

$$S_T = S_1 + S_2 + \cdots + S_K \tag{2.39}$$

where K is a geometric random variable with

$$Pr(K = k) = p^{k-1}(1 - p), \quad (k \geq 1) \tag{2.40}$$

Since the probability generating function $K(z)$ of K is $[(1-p)z]/(1-pz)$, we have the Laplace transform of the probability density function of S_T as

$$S_T^*(\theta) = \frac{(1-p)S^*(\theta)}{1 - pS^*(\theta)} \tag{2.41}$$

The mean becomes

$$E(S_T) = \frac{E(S)}{1-p} \tag{2.42}$$

and the second moment becomes

$$E(S_T^2) = \frac{(1-p)E(S^2) + 2pE^2(S)}{(1-p)^2} \tag{2.43}$$

Let $\Pi_{FB}^*(z, \theta)$ be the joint transform of the number of cells in the system and the remaining total processing time in an ordinary $M/G/1$ system with Bernoulli feedback (we will denote the system as $M/G/1/FB$). Then, it is known that (Takagi [15])

$$\Pi_{FB}^*(z, \theta) = \frac{\lambda(1 - \rho_T)z(1 - z)[S_T^*(\lambda - \lambda z) - S_T^*(\theta)]}{(\theta - \lambda + \lambda z)[S_T^*(\lambda - \lambda z) - z]} \tag{2.44}$$

where $\rho_T = \lambda E(S_T) = \lambda E(S)/(1-p)$. Thus the mean remaining total processing time when the entering cell sees one cell in the ordinary

$M/G/1/FB$ system becomes

$$E(T_{R,1,M/G/1/FB}) = \frac{-(\partial/\partial\theta)(\partial/\partial z)\Pi^*_{FB}(z,\theta)\big|_{z=0,\theta=0}}{P_{1,M/G/1/FB}}$$
$$= \frac{(1-\rho_T)}{\lambda P_{1,M/G/1/FB}}\left[1+\frac{E(S_T)}{S^*_T(\lambda)}\right], \tag{2.45}$$

where $P_{1,M/G/1/FB}$ is the probability that an arriving cell finds one cell in the system in the ordinary $M/G/1/FB$ queue.

Returning to our system, let $E(T_{R,j})$ be the mean remaining total processing time of the in-service cell under the condition that the entering cell finds j cells in the system. To obtain $E(T_{R,1})$, we apply the result of ordinary $M/G/1$ queueing system with infinite capacity. Our justification for doing so is that the loss probability in our system is extremely low. Thus using $Q^1_2 = P_1/[1 - P_{\text{Loss}\,1}]$ in place of $P_{1,M/G/1/FB}$ in (2.45), we can approximate $E(T_{R,1})$ by $\hat{E}(T_{R,1})$ where

$$\hat{E}(T_{R,1}) = \frac{(1-\rho_T)}{\lambda Q^1_2}\left[1+\frac{E(S_T)}{S^*_T(\lambda)}\right]. \tag{2.46}$$

$E(T_{R,n,M/G/1/FB})$ for $n \geq 2$ is hard to obtain because we need to differentiate (2.44) n times. Thus in these cases, we are going to use the remaining total processing time irrespective of the number of cells seen at an arrival epoch. Let S_{T+} be the steady-state remaining total processing time that an arbitrary arriving test cell sees irrespective of the number of cells in an ordinary $M/G/1/FB$ system. Then we have (Wolff [19])

$$E(S_{T+}) = \sum_{n=1}^{\infty} P_{n,M/G/1/FB} \cdot E(T_{R,n,M/G/1/FB}) = \frac{E(S_T^2)}{2E(S_T)}$$
$$= \frac{(1-p)E(S^2) + 2pE^2(S)}{2(1-p)E(S)} \tag{2.47a}$$

or

$$\sum_{n=2}^{\infty} P_{n,M/G/1/FB} \cdot E(T_{R,n,M/G/1/FB})$$
$$= E(S_{T+}) - P_{1,M/G/1/FB}E(T_{R,1,M/G/1/FB}). \tag{2.47b}$$

Thus we have

$$E(T_{R,n,M/G/1/FB}) = E(T_{R,M/G/1/FB}) = \frac{E(S_{T+}) - P_{1,M/G/1/FB}E(T_{R,1,M/G/1/FB})}{\sum_{n=2}^{\infty} P_{n,M/G/1/FB}}. \tag{2.48}$$

Thus for $n \geq 2$, we use the following quantity as the approximate mean remaining total processing time averaged over all cells:

$$\hat{E}(T_{R,n}) = \hat{E}(T_R) = \frac{E(S_{T+}) - Q_2^1 \hat{E}(T_{R,1})}{1 - Q_1^1 - Q_2^1}, \quad (n \geq 2). \tag{2.49}$$

In summary, we use the following quantity as the approximate mean total remaining processing time:

$$\hat{E}(T_{R,n}) = \begin{cases} \dfrac{(1-\rho_T)}{\lambda Q_2^1}\left[1 + \dfrac{E(S_T)}{S_T^*(\lambda)}\right] & (n = 1), \\ \dfrac{E(S_{T+}) - Q_2^1 \hat{E}(T_{R,1})}{1 - Q_1^1 - Q_2^1} & (n \geq 2). \end{cases} \tag{2.50}$$

We, then, have the approximate formula for mean waiting time of an arbitrary high-priority cell as

$$W_q^{(1)} = Q_2^1 \hat{E}(T_{R,1}) + \sum_{n=3}^{K+1} Q_n^1 \cdot \{\hat{E}(T_{R,n}) + (n-2)E(S_T)\}. \tag{2.51}$$

The mean system sojourn time $W^{(1)}$ of high-priority cells can be obtained from

$$W^{(1)} = E(S_T) + W_q^{(1)}. \tag{2.52}$$

Similarly, we obtain the approximate mean waiting time and sojourn time of a low-priority cell as

$$W_q^{(2)} = Q_2^2 \hat{E}(T_{R,1}) + \sum_{n=3}^{T+1} Q_n^2 \cdot \{\hat{E}(T_{R,n}) + (n-2)E(S_T)\}, \tag{2.53}$$

$$W^{(2)} = E(S_T) + W_q^{(2)}. \tag{2.54}$$

3. PERFORMANCE ANALYSIS

In this section we analyze the performance of the system.

3.1. Changing the Feedback Probability p

We first analyze the effect of the feedback probability p on the loss probabilities and mean system sojourn time. We see in Fig. 2 that as p increases $P_{\text{Loss}\,1}$ and $P_{\text{Loss}\,2}$ increase. For this particular example, loss probabilities begin to increase sharply around at $p=0.1$. Figure 3 shows that as p increases mean system sojourn time of both types increases.

3.2. Changing the Load Ratio λ_1/λ

Figure 4 shows the effect of the load ratio λ_1/λ on the mean system sojourn time for different p. As we can see in the figure, p has more effect on the mean system sojourn time than the load ratio λ_1/λ does. This tells us that controlling p is more effective than controlling the load ratio.

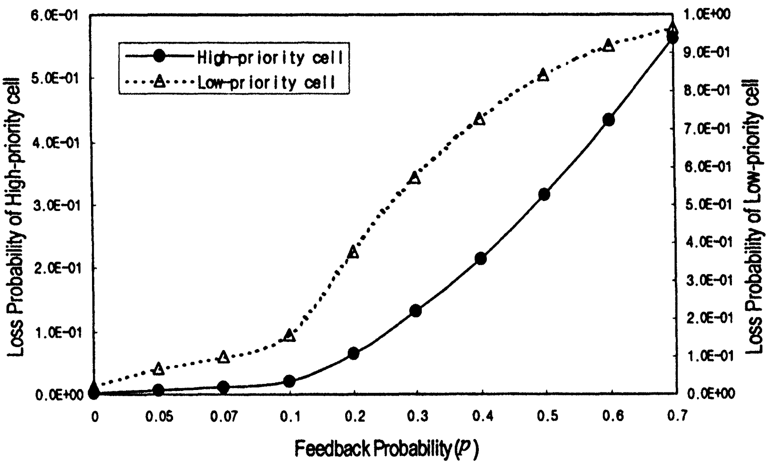

FIGURE 2 Change of loss probabilities as p varies ($a=0.95$, $\lambda_1/\lambda=0.7$): $\lambda=5$, $E(S)=0.19$, $\text{Var}(S)=0$, $T=18$, $K=20$.

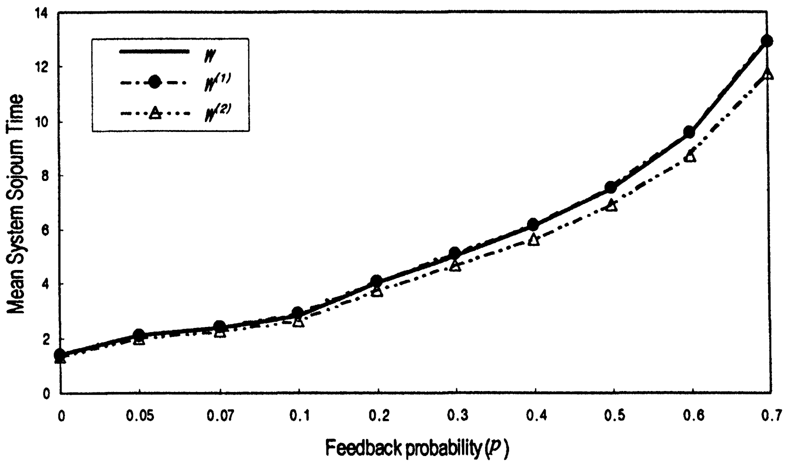

FIGURE 3 Mean system sojourn times as p varies ($a=0.95$, $\lambda_1/\lambda=0.7$): $\lambda=5$, $E(S)=0.19$, $\text{Var}(S)=0$, $T=18$, $K=20$.

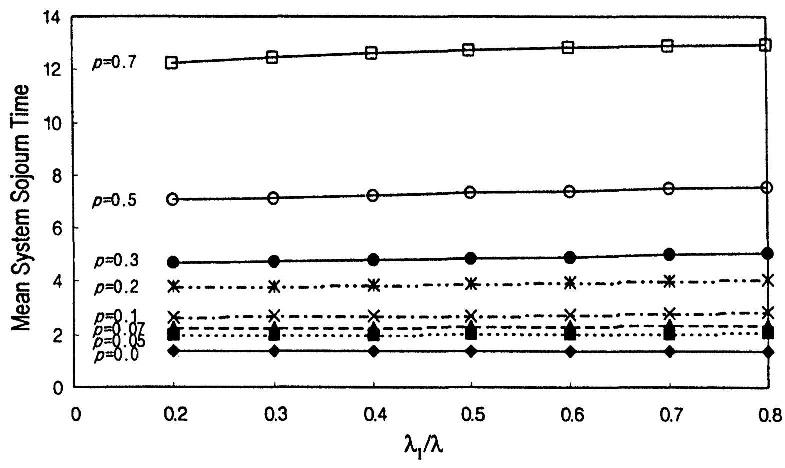

FIGURE 4 Mean system sojourn times as λ_1/λ varies for different feedback probabilities p ($a=0.95$): $\lambda=5$, $E(S)=0.19$, $\text{Var}(S)=0$, $T=18$, $K=204$.

4. OPTIMAL THRESHOLD T^* AND BUFFER SIZE K^*

In this section we propose an algorithm to determine the optimal threshold T^* and buffer size K^* that satisfy a given QoS. Let $LIMIT_1$

and $LIMIT_2$ be the QoSs for high- and low-priority cells respectively. We define (T^*, K^*) as the minimum values of T and K that satisfies the QoSs simultaneously. Then we have

$$(T^*, K^*) = \min\{(T, K) | P_{\text{Loss}\,1} \leq LIMIT_1 \text{ and } P_{\text{Loss}\,2} \leq LIMIT_2\}. \tag{4.1}$$

Since the functional form of P_n is not known, it is impossible to find optimal T^* and K^* analytically.

4.1. The Algorithm

The following algorithm is developed based on (4.1).

(*Algorithm for finding the optimal threshold* T^* *and queue capacity* K^*)

⟨step 1⟩ Set $K = 1$.
⟨step 2⟩ Set $T = 1$.
⟨step 3⟩ Compute $P_{\text{Loss}\,1}$ and $P_{\text{Loss}\,2}$.
⟨step 4⟩ If $P_{\text{Loss}\,1} \leq LIMIT_1$ and $P_{\text{Loss}\,2} \leq LIMIT_2$
  Set $I_1 = T$, $I_2 = K$, go to ⟨step 6⟩.
  Otherwise, go to ⟨step 5⟩.
⟨step 5⟩ Set $T = T + 1$,
  If $T > K$
  Set $K = K + 1$, go to ⟨step 2⟩.
  Otherwise, go to ⟨step 3⟩.
⟨step 6⟩ Stop. $(T^*, K^*) = (I_1, I_2)$.

4.2. Numerical Example

Suppose we have $\lambda = 1$, $p = 0.01$, $E(S) = 0.9$, $\text{Var}(S) = 0$, $LIMIT_1 = 10^{-5}$ and $LIMIT_2 = 10^{-3}$. We find the optimal pair of the threshold and the queue capacity (T^*, K^*) as the load ratio λ_1/λ varies. Figure 5 shows the optimal values (T^*, K^*) for various load ratios. We see that (T^*, K^*) also increases as the ratio λ_1/λ increases. For this particular example, partial buffer sharing does not make sense if the load ratio is greater than 0.7.

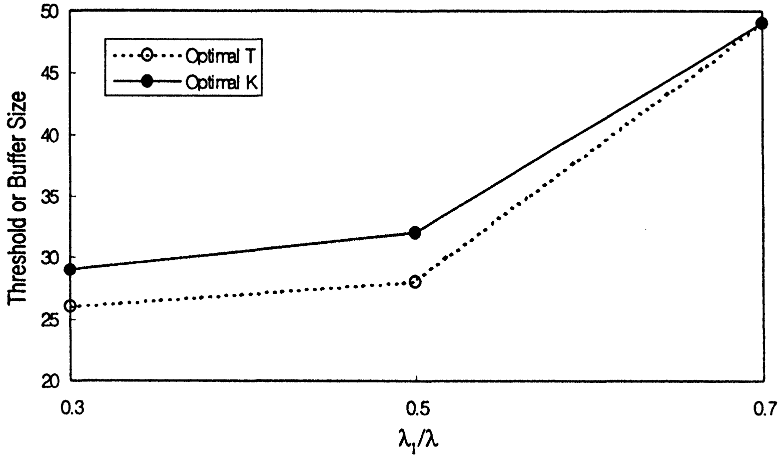

FIGURE 5 Optimal threshold and queue capacity: $a=0.9$, $LIMIT_1=10^{-5}$, $LIMIT_2=10^{-3}$, $\lambda=1$, $p=0.01$, $E(S)=0.9$, $\mathrm{Var}(S)=0$.

5. RESEARCH SUMMARY

We analyzed the partial buffer sharing with error control using $M_1, M_2/G/1/K+1$ queue with threshold and instantaneous Bernoulli feedback. We developed a recursive algorithm to find the loss probabilities. We also obtained the approximate mean sojourn times of each class of cells. We finally proposed an algorithm that determines the optimal threshold T^* and buffer size K^* under a given QoS.

References

[1] I.F. Akyildiz and X. Cheng, "Analysis of a finite buffer queue with different scheduling and push-out schemes", *Performance Evaluation*, **19**, 317–340, 1994.

[2] J.J. Bae and T. Suda, "Survey of traffic control schemes and protocols in ATM networks", *Proceedings of the IEEE*, **79**(2), 170–189, 1991.

[3] CCITT WP XV III/8 Draft Recommendation I.371, "Traffic Control and Resource Management in B-ISDN", Melbourne, 1991.

[4] C.G. Chang and H.H. Tan, "Queueing analysis of explicit policy assignment push-out buffer sharing schemes for ATM networks", *Proceedings of IEEE INFOCOM'94*.

[5] J. Garcia and O. Casals, "Stochastic models of space priority mechanisms with Markovian arrival processes", *Annals of Operations Research*, **35**, 271–296, 1992.

[6] U.C. Gupta and T.S.S. Srini Vasa Rao, "Computing steady state probabilities in $\lambda(n)/G/1/K$ queue", Working paper, Department of Mathematics, Indian Institute of Technology, June, 1993.

[7] U.C. Gupta and T.S.S. Srini Vasa Rao, "A recursive method to compute the steady state probabilities of the machine interference model: $(M/G/1)/K$", *Computers Opns. Res.*, **21**(6), 597–605, 1994.

[8] G. Hebuterne and A. Gravey, "A space priority queueing mechanism for multiplexing ATM channels", ITC Specialist Seminar, Adelaide, Paper 7.4, September 1989.

[9] N.K. Jaiswal, *Priority Queue*, Academic Press, New York, 1968.

[10] H. Kröner, "Comparative performance study of space priority mechanisms for ATM channels", *Proceedings of IEEE INFOCOM'90*, 1136–1143.

[11] H.W. Lee and B.Y. Ahn, "A queueing model for optimal control of partial bufffer sharing in ATM", *Computers Opns. Res.*, **25**(2), 113–126, 1998.

[12] M. Murata, Y. Kikuchi, H. Inai and H. Miyahara, "Performance comparison of error recovery schemes for high bandwidth-delay product networks", ITC 14, 1994.

[13] D.W. Petr and V.S. Frost, "Optimal threshold-based discarding for queue overload control", *Proceedings of the 7th ITC seminar*, Morristown, Paper 10.5, October 1990.

[14] S. Suri, D. Tipper and G. Meempat, "A comparative evaluation of space priority strategies in ATM networks", *Proceedings of IEEE INFOCOM '94*.

[15] H. Takagi, *Queueing Analysis: A Foundation of Performance Evaluation, Vol. I, Vacation and Priority Systems, Part I*, North-Holland, 1991.

[16] H. Takagi, *Queueing Analysis: A Foundation of Performance Evaluation, Vol. II, Finite Systems*, North-Holland, 1991.

[17] H. Takagi, *Queueing Analysis: A Foundation of Performance Evaluation, Vol. III, Discrete Systems*, North-Holland, 1991.

[18] J. Walrand, *Communication Networks: A First Course*, Richard D. Irwin, Inc. and Aksen Associates, Inc., 1991.

[19] R.W. Wolff, "Poisson Arrivals See Time Average", *Opns. Res.*, **30**(2), 223–231, 1986.

[20] P. Yang, "A unified algorithm for computing the stationary queue length distribution in $M(k)/G/1/N$ and $GI/M(k)/N$ queues", *Queueing Systems*, **17**, 383–401, 1994.