Journal of the Operations Research
Society of Japan
Vol. 40, No. 1, March 1997

CELL-LOSS-RATIO ANALYSIS WITH INSUFFICIENT KNOWLEDGE OF TRAFFIC CHARACTERISTICS

Shigeo Shioda   Hiroshi Saito
NTT Multimedia Networks Laboratories

(Received January 8, 1996; Final October 24, 1996)

Abstract This paper evaluates cell-loss ratio in an output buffer of an ATM node based on the observed relative frequency of the number of cell arrivals during a fixed interval. The central issue of this evaluation problem is that the unique evaluation result of cell-loss ratio cannot be derived because the relative frequency of the number of cell arrivals does not completely describe the traffic characteristics of cell streams. Thus, this paper focuses on the derivation of the worst-case performance, that is, the cell-loss-ratio upper bound. Two cell-loss-ratio upper bounds are derived – one for when the cell streams are stationary, and one for when cell streams are stationary and ergodic. Numerical results show that cell streams with the same relative frequency of the number of cell arrivals can have quite different cell-loss ratios, and that the derived formula gives the actual upper bounds.

1. Introduction

In asynchronous transfer mode (ATM) networks, the anticipated traffic characteristics of the connection are specified by a reduced set of parameters, called a source traffic descriptor (STD), with quality-of-service (QOS) requirements at the connection set-up phase. The network assigns network resources for the connection on the basis of the STD values and QOS requirements. If there are not enough network resources, the connection is rejected. If the connection is accepted, actual traffic characteristics are checked for conformity with the source traffic descriptor specified at connection set-up by monitoring the cell stream of the connection. To simplify traffic specification and monitoring, the STDs must be observable and manageable through conformance testing. Algorithmic or rule-based STDs currently being discussed in the ITU satisfy these requirements. Such STDs, however, do not simplify connection admission control and resource management. For instance, algorithmic STDs allow a wide variety of traffic patterns to enter the network, and the network must accept these patterns without violating the QOS requirements.

Connection admission control (CAC) for sources characterized by algorithmic STDs has motivated the study of QOS evaluation for the "worst case" pattern of all possible patterns that are compliant with the STD. Doshi [6] formally addressed this problem for the case where users comply with their own leaky-bucket-based STD and showed that so-called on/off patterns, which are generally believed to be the worst traffic pattern, are not always the worst. Based on analysis of queue-length distributions, Yamanaka et al. [18] gave a nonintuitive example in which on/off sources may not be the worst. Of all possible output processes from a leaky bucket, Lee [12] identified the worst traffic pattern that maximizes the average waiting time.

In this paper we consider this problem from a slightly different viewpoint: the traffic characteristics of the source are described by the relative frequency of the number of cells arriving during a fixed interval, instead of by the STD. This relative frequency can be constructed from the STD values; it can also be estimated by direct measurement. A cell-loss-ratio (CLR) evaluation formula based on the relative frequency might be used as the basis of adaptive CAC, in which connection acceptance is based on the STD values and the cell-arrival-process measurement.

The relative frequency of the number of cells arriving during a fixed interval is not complete information about the traffic characteristics: the relative frequency of the number of cells is just steady state statistics, and higher-order statistics like autocorrelation functions are further needed to describe all the traffic characteristics precisely. System performance evaluation using insufficient information about traffic characteristics does not give unique evaluation results and, for conservative evaluation of system performance, the worst case performance should be identified. This paper attempts to derive the upper bounds of the cell-loss ratio based only on the relative frequency of the number of arrival cells. (To simplify the discussions, however, some basic traffic characteristics, like a stationary state or ergodicity, are further assumed.)

Note that the relative frequency of the number of cells might depend on the observation conditions even if the underlying traffic characteristic is unique. More presicely, the relative frequency, which is a time-averaged quantity, might be a random variable on some probability space (see Section 2), and thus it has sample-path dependency. Furthermore, the observed CLR itself has sample-path dependency even if the underlying traffic characteristic is unique because it is a time-averaged quantity. This requires some mathematical formulation that enables us to evaluate the sample-path-wise CLR from the relative frequency observed for each sample path. Such a formulation is very reasonable because, in principle, the QOS of an ATM network (e.g. cell-loss ratio or queueing delay) should be guaranteed for each user. The QOS should thus be defined by the long-term average quality for the sample path, although traditional queueing analyses generally evaluate QOS values based on their ensemble averages, assuming some ergodicity. The ergodic assumption, however, may not be valid in an ATM environment because some users are likely to transfer cells in periodic patterns. In this paper, we therefore propose the CLR upper bounds that hold for each sample path.

Our results include and generalize those of related works regarding the cell-loss-ratio upper bounds [13, 14, 15]: for instance, we derive the sample-path version of the formula in Ref. [13]. We also complete the proof of the cell-loss-ratio upper bound used in [14] and give the sufficient condition under which it holds. After introducing the cell arrival process in Section 2, we derive two different cell-loss-ratio upper bounds in Section 3, without assuming that the traffic source is ergodic. In Section 4, we derive a different upper bound under an ergodic assumption and show that the ergodic assumption makes the bound tighter. In Section 5, we present numerical examples that confirm the tightness of the derived upper bound. In Section 6, the "effective bandwidth" for a bufferless model is derived based on the proposed cell-loss-ratio upper bound. The derived effective bandwidth is an extended version of that proposed by Kelly [9, 3]. Finally, in Section 7, we summarize the key points of this work.

2. Cell-arrival process

The time axis is assumed to be divided into slots with lengths equal to cell transmission time L/C, where L is cell length and C is virtual path bandwidth. A cell arrives at the beginning of a slot and is transmitted at the end of a slot. Each slot is sequentially numbered such that the 0th slot ends (the 1st slot starts) at the time origin (Fig. 1). Let $a_i^{(1)}$ be the

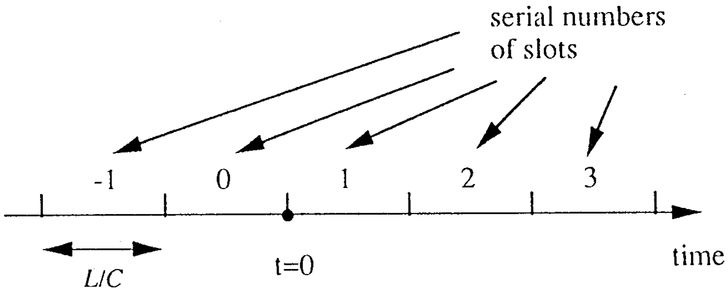

Figure 1. Slot numbering on the time axis.

number of cells arriving during the i-th slot; $\{a_i^{(1)}\}_{i=-\infty}^{\infty}$ is assumed to be a stochastic process defined on the probability space $(\Omega, \mathcal{F}, P)$. In $(\Omega, \mathcal{F})$, the time-shift operator $\{\theta_i\}$, $i \in \boldsymbol{N}$, is as defined in [1], and P is θ_i-invariant for all i. Suppose that the stochastic process $\{a_i^{(1)}\}_{i=-\infty}^{\infty}$ is stationary, that is [1],

$$a_i^{(1)}(\theta_j \circ \omega) = a_{i+j}^{(1)}(\omega) \quad \text{for all } \omega \in \Omega. \tag{2.1}$$

We construct the definition period sequence on a time axis (Fig. 2): we denote a period

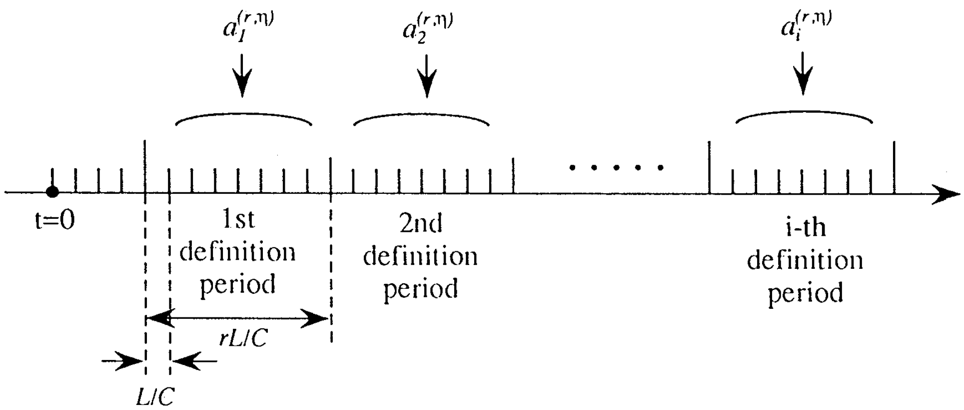

Figure 2. Definition period sequence.

of length r ($r \in \boldsymbol{N}$) slots as a definition period and give each definition period a sequential number. The 1st definition period starts at the η-th slot ($1 \le \eta \le r$). (For example, in Fig. 2, $\eta = 5$.) We introduce two supplementary stochastic processes, $\{x_i^r\}_{i=-\infty}^{\infty}$ and $\{a_n^{(r,\eta)}\}_{n=-\infty}^{\infty}$:

$$x_i^r \stackrel{\text{def}}{=} \sum_{j=0}^{r-1} a_{i+j}^{(1)}; \qquad \text{and} \qquad a_n^{(r,\eta)} \stackrel{\text{def}}{=} x_{r(n-1)+\eta}^r. \tag{2.2}$$

Here, $a_n^{(r,\eta)}$ effectively corresponds to the number of cells arriving during the n-th definition period, whose length is equal to rL/C. Obviously, $\{x_i^r\}_{i=-\infty}^{\infty}$ and $\{a_n^{(r,\eta)}\}_{n=-\infty}^{\infty}$ are stationary. Let $\mathcal{F}_x$ be the σ-field generated by $\{x_i^r\}_{i=-\infty}^{\infty}$ and $\mathcal{G}_r$ be the set of $A \in \mathcal{F}_x$ defined by

$$\mathcal{G}_r = \{A : \mathbf{1}_{\theta_r A} = \mathbf{1}_A \quad P\text{–a.s.}\}. \tag{2.3}$$

The following lemma guarantees that the relative frequency of $\{a_n^{(r,\eta)}\}_{n=-\infty}^{\infty}$ can be defined for each sample path. For proof of this, please refer to a book on ergodic theory, e.g. [2].

Lemma 2.1 (Ergodic Theorem) *When* $\{a_i^{(1)}\}_{i=-\infty}^{\infty}$ *is stationary,*

(2.4) $$p(k;r,\eta) \stackrel{\text{def}}{=} \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}(a_n^{(r,\eta)} = k) = E[\mathbf{1}(a_1^{(r,\eta)} = k) \mid \mathcal{G}_r] \quad P\text{–a.s.}$$

The above means that the relative frequency, $p(k;r,\eta)$ is a random variable measurable on $\mathcal{G}_r$.

3. Cell-loss-ratio upper bounds

An ATM node consists of output buffers and a high speed switch (Fig. 3), and the latter is nearly nonblocking because of advanced LSI technology. We therefore focus on the cell losses in an output buffer of an ATM node. Since the cell is usually served on a first-in-first-out (FIFO) basis and the cell length is fixed (53 bytes), an output buffer can be modeled by a G/D/1/K+1 system with a FIFO service discipline, where K is the size of output buffer [16].

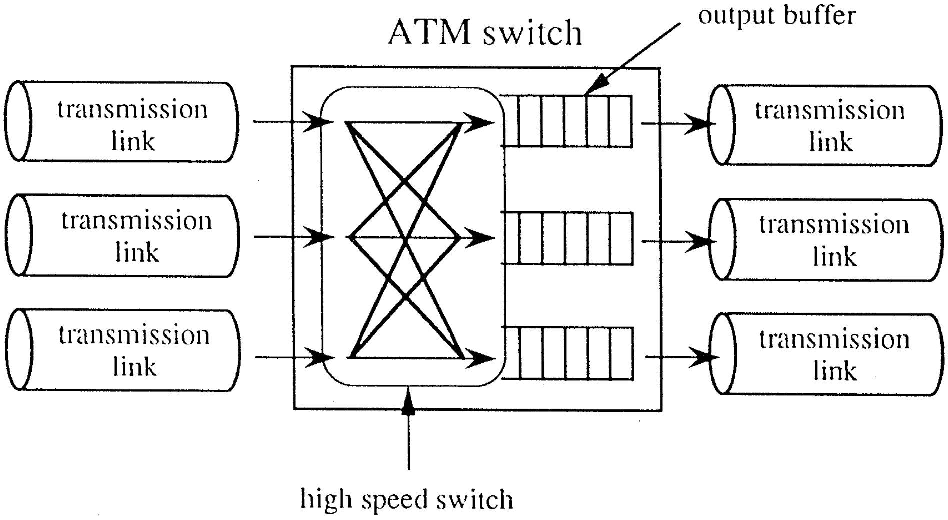

Figure 3. ATM switch.

A busy period of this queueing model in which at least one cell is lost is denoted here as the lossy busy period (LBP). The LBP structure of this system is shown in Fig. 4. The l-th LBP starts at the $\{(n_s(l)-1)r + t_s(l) + (\eta-1)\}$-th slot (within the $n_s(l)$-th definition period), and the last cell loss in the l-th LBP occurs at the $\{(n_e(l)-1)r + t_e(l) + (\eta-1)\}$-th slot (within the $n_e(l)$-th definition period). Let $b(l)$ be the number of lost cells within the l-th LBP. The cell-loss ratio (CLR) is defined in terms of time average:

(3.1) $$CLR \stackrel{\text{def}}{=} \lim_{N\to\infty} \frac{\text{number of lost cells during } [0, rN]}{\text{number of cells arriving during } [0, rN]}$$

where $[x, y]$ means the range from the x-th to y-th slots. Let $L(N)$ be the serial number of the latest LBP among those which have started by the end of the N-th definition period. Since

(3.2) $$\text{number of lost cells during } [0, rN] \le \sum_{l=1}^{L(N)} b(l),$$

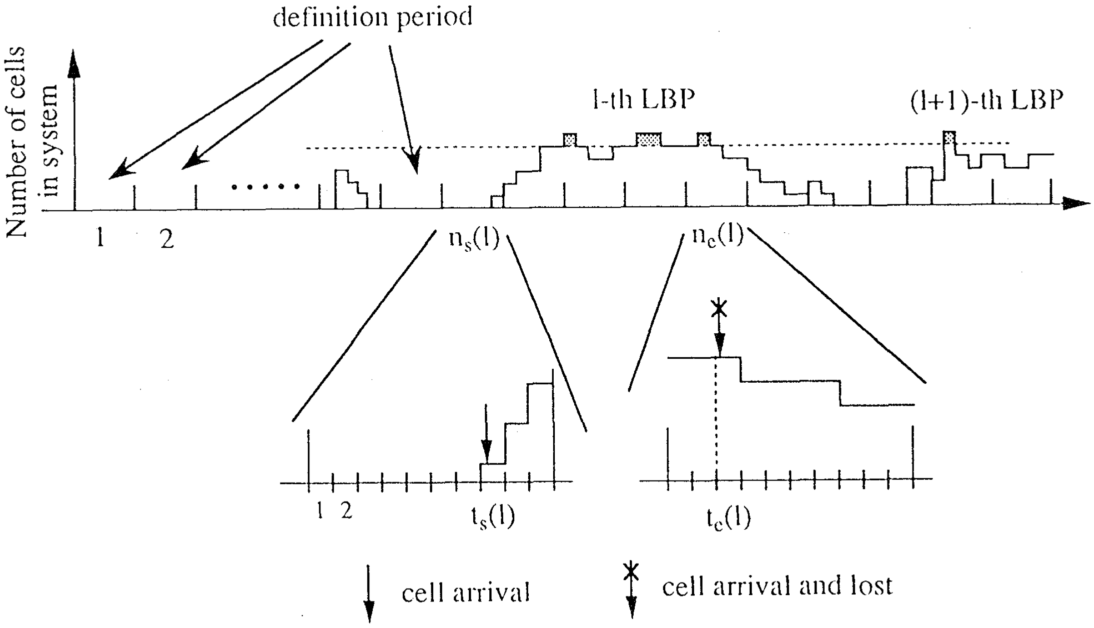

Figure 4. Example of lossy busy period (LBP) in the definition period sequence.

we have

$$CLR \leq \lim_{N\to\infty} \frac{\sum_{l=1}^{L(N)} b(l)}{\sum_{n=1}^{N} a_n^{(r,\eta)}}. \tag{3.3}$$

The values of parameters $t_s(l), t_e(l), n_s(l)$, and $n_e(l)$ change when the definition-period sequence is shifted along the time axis. If necessary, we use notation like $t_s(l;\eta)$ or $n_s(l;\eta)$ to express how the definition-period sequence is defined on the time axis.

Theorem 3.1 *For all η,*

$$CLR \leq \frac{\sum_k [k - \min\{r, (K+2)/2\}]^+ p(k;r,\eta)}{\sum_k kp(k;r,\eta)} \qquad P\text{–a.s.}. \tag{3.4}$$

Proof: The number of lost cells within the l-th LBP, $b(l)$, is equal to the number of arriving cells during $[(n_s(l)-1)r + t_s(l) + (\eta-1), (n_e(l)-1)r + t_e(l) + (\eta-1)]$ minus the number of cells transmitted during $[(n_s(l)-1)r + t_s(l) + (\eta-1), (n_e(l)-1)r + t_e(l) + (\eta-1)]$ and waiting in the buffer at the end of the $\{(n_e(l)-1)r + t_e(l) + (\eta-1)\}$-th slot. When $n_s(l) = n_e(l)$, we have

$$b(l) = a_{n_s(l)}^{(l,r,\eta)} - \{(t_e(l) - t_s(l)) + K + 1\} \leq a_{n_s(l)}^{(l,r,\eta)} - (K+1). \tag{3.5}$$

Here, $a_{n_s(l)}^{(l,r,\eta)}$ denotes the number of cells arriving during the $n_s(l)$-th definition period within the l-th LBP.

When $n_s(l) < n_e(l)$,

$$\begin{aligned} b(l) = (a_{n_s(l)}^{(l,r,\eta)} + a_{n_s(l)+1}^{(r,\eta)} + \cdots + a_{n_e(l)-1}^{(r,\eta)} + a_{n_e(l)}^{(l,r,\eta)}) \\ -\{(n_e(l) - n_s(l))r + (t_e(l) - t_s(l)) + K + 1\}, \end{aligned} \tag{3.6}$$

where $a_{n_e(l)}^{(l,r,\eta)}$ is the number arriving during the $n_e(l)$-th definition period until the $t_e(l)$-th slot. Equation (3.6) can be simplified as follows:

$$(3.7)\quad \begin{aligned} b(l) &= (a_{n_s(l)}^{(l,r,\eta)} + a_{n_s(l)+1}^{(r,\eta)} + \cdots + a_{n_e(l)-1}^{(r,\eta)} + a_{n_e(l)}^{(l,r,\eta)}) + (t_s(l) - t_e(l) - r) \\ &\qquad -(n_e(l) - n_s(l) - 1)r - (K+1) \\ &\le (a_{n_s(l)}^{(l,r,\eta)} + a_{n_s(l)+1}^{(r,\eta)} + \cdots + a_{n_e(l)-1}^{(r,\eta)} + a_{n_e(l)}^{(l,r,\eta)}) - (n_e(l) - n_s(l) - 1)r - (K+2) \\ &\le (a_{n_s(l)}^{(l,r,\eta)} + a_{n_s(l)+1}^{(r,\eta)} + \cdots + a_{n_e(l)-1}^{(r,\eta)} + a_{n_e(l)}^{(l,r,\eta)}) - (n_e(l) - n_s(l) + 1)\gamma_1(r) \\ &= (a_{n_s(l)}^{(l,r,\eta)} - \gamma_1(r)) + \sum_{n=n_s(l)+1}^{n_e(l)-1} (a_n^{(r,\eta)} - \gamma_1(r)) + (a_{n_e(l)}^{(l,r,\eta)} - \gamma_1(r)), \end{aligned}$$

where $\gamma_1(r) = \min\{r, (K+2)/2\}$. For both cases, $b(l)$ is upper-bounded by Eq. (3.7), and thus

$$(3.8)\quad \begin{aligned} CLR &\le \lim_{N\to\infty} \frac{\sum_{l=1}^{L(N)} \left\{(a_{n_s(l)}^{(l,r,\eta)} - \gamma_1(r)) + \sum_{n=n_s(l)+1}^{n_e(l)-1} (a_n^{(r,\eta)} - \gamma_1(r)) + (a_{n_e(l)}^{(l,r,\eta)} - \gamma_1(r))\right\}}{\sum_{n=1}^{N} a_n^{(r,\eta)}} \\ &\le \lim_{N\to\infty} \frac{\sum_{n=1}^{n_e(L(N))} [a_n^{(r,\eta)} - \gamma_1(r)]^+}{\sum_{n=1}^{N} a_n^{(r,\eta)}} \\ &= \lim_{N\to\infty} \frac{\sum_k \sum_{n=1}^{N} [k - \gamma_1(r)]^+ \mathbf{1}(a_n^{(r,\eta)} = k)}{\sum_k \sum_{n=1}^{N} k\mathbf{1}(a_n^{(r,\eta)} = k)}. \end{aligned}$$

The final version of the right-hand side converges with probability 1 and can be rewritten with the frequency distribution $\{p(k;r,\eta)\}$ defined by Eq. (2.4). Consequently,

$$(3.9)\qquad CLR \le \frac{\sum_k [k - \min\{r, (K+2)/2\}]^+ p(k;r,\eta)}{\sum_k kp(k;r,\eta)} \quad P\text{–a.s.} \qquad \blacksquare$$

For use in a later section, we here prove the following corollary.

Corollary 3.1 *For all* η

$$(3.10)\qquad CLR \le \frac{\sum_k [k - \min\{r, K+1\}]^+ p(k;r,\eta)}{\sum_k kp(k;r,\eta)} + q(r,\eta) \quad P\text{–a.s.},$$

where

$$(3.11)\qquad q(r,\eta) \stackrel{\text{def}}{=} \frac{\lim_{L\to\infty} \frac{1}{n_e(L)} \sum_{l=1}^{L} (t_s(l;\eta) - t_e(l;\eta))}{\sum_k kp(k;r,\eta)}.$$

Proof: Equation (3.6) can be transformed into

$$
\begin{aligned}
(3.12)\quad b(l) &\le (a_{n_s(l)}^{(l,r,\eta)} + a_{n_s(l)+1}^{(r,\eta)} + \cdots + a_{n_e(l)-1}^{(r,\eta)} + a_{n_e(l)}^{(l,r,\eta)}) + (t_s(l) - t_e(l)) \\
&\qquad - (n_e(l) - n_s(l))r - (K+1) \\
&\le (a_{n_s(l)}^{(l,r,\eta)} + a_{n_s(l)+1}^{(r,\eta)} + \cdots + a_{n_e(l)-1}^{(r,\eta)} + a_{n_e(l)}^{(l,r,\eta)}) + (t_s(l) - t_e(l)) \\
&\qquad - (n_e(l) - n_s(l) + 1)\gamma_2(r) \\
&= (a_{n_s(l)}^{(l,r,\eta)} - \gamma_2(r)) + \sum_{n=n_s(l)+1}^{n_e(l)-1} (a_n^{(r,\eta)} - \gamma_2(r)) \\
&\qquad + (a_{n_e(l)}^{(l,r,\eta)} - \gamma_2(r)) + (t_s(l) - t_e(l)),
\end{aligned}
$$

where $\gamma_2(r) = \min\{r, K+1\}$. Using Eq. (3.12), we get

$$
\begin{aligned}
CLR &\le \lim_{N\to\infty} \frac{\sum_{l=1}^{L(N)} \{(a_{n_s(l)}^{(l,r,\eta)} - \gamma_2(r)) + \sum_{n=n_s(l)+1}^{n_e(l)-1} (a_n^{(r,\eta)} - \gamma_2(r)) + (a_{n_e(l)}^{(l,r,\eta)} - \gamma_2(r))\}}{\sum_{n=1}^{N} a_n^{(r,\eta)}} \\
&\qquad + \lim_{N\to\infty} \frac{\sum_{l=1}^{L(N)} (t_s(l) - t_e(l))}{\sum_{n=1}^{N} a_n^{(r,\eta)}} \\
(3.13)\qquad &\le \lim_{N\to\infty} \frac{\sum_{n=1}^{n_e(L(N))} [a_n^{(r,\eta)} - \gamma_2(r)]^+}{\sum_{n=1}^{N} a_n^{(r,\eta)}} + \lim_{N\to\infty} \frac{\sum_{l=1}^{L(N)} (t_s(l) - t_e(l))}{\sum_{n=1}^{N} a_n^{(r,\eta)}} \\
&= \frac{\sum_k [k - \gamma_2(r)]^+ p(k; r, \eta)}{\sum_k k p(k; r, \eta)} + \frac{\lim_{N\to\infty} \frac{1}{n_e(L)} \sum_{l=1}^{L} (t_s(l;\eta) - t_e(l;\eta))}{\sum_k k p(k; r, \eta)}. \qquad \blacksquare
\end{aligned}
$$

4. Cell-loss-ratio upper bound under an ergodic assumption

In this section, we will derive the cell-loss-ratio upper bound under an ergodic assumption. First, we introduce the concept of ergodicity in probability theory [2]:

Definition 4.1 *The operator θ_1 is said to be ergodic for $(\Omega, \mathcal{F}_x, P)$ if all sets $A \in \mathcal{F}_x$ such that $\mathbf{1}_A = \mathbf{1}_{\theta_1 A}$ with probability 1 satisfy $P(A) = 0$ or 1.*

Lemma 4.1

$$(4.1)\qquad \min_\eta q(r, \eta) \le 0.$$

Proof: We have

$$(4.2)\quad \sum_{\eta=1}^{r} \lim_{L\to\infty} \frac{1}{n_e(L)} \sum_{l=1}^{L} (t_s(l;\eta) - t_e(l;\eta)) = \lim_{L\to\infty} \frac{1}{n_e(L)} \sum_{l=1}^{L} \Big(\sum_{\eta=1}^{r} t_s(l;\eta) - \sum_{\eta=1}^{r} t_e(l;\eta)\Big) = 0.$$

To derive the final equality, we use the following relationship:

$$(4.3)\qquad \sum_{\eta=1}^{r} t_s(l;\eta) = \sum_{\eta=1}^{r} t_e(l;\eta) = (r+1)r/2.$$

Thus, we have

$$\min_{\eta} q(r,\eta) = \min_{\eta}\left\{\lim_{L\to\infty}\frac{1}{n_{e(L)}}\sum_{l=1}^{L}(t_s(l;\eta)-t_e(l;\eta))/\sum_k kp(k;r,\eta)\right\} \tag{4.4}$$
$$\leq \frac{1}{r}\left\{\sum_{\eta=1}^{r}\lim_{L\to\infty}\frac{1}{n_{e(L)}}\sum_{n=1}^{L}(t_s(l;\eta)-t_e(l;\eta))/\sum_k kp(k;r,\eta)\right\}$$
$$= \frac{1}{r\sum_k kp(k;r,\eta)}\lim_{L\to\infty}\frac{1}{n_{e(L)}}\sum_{l=1}^{L}\sum_{\eta=1}^{r}\{t_s(l;\eta)-t_e(l;\eta)\}$$
$$= 0. \qquad \blacksquare$$

Lemma 4.2 *Suppose that time-shift operator θ_r is ergodic for $(\Omega,\mathcal{F}x,P)$. Then, $\{p(k;r,1)\}_{k=0}^{\infty}=\{p(k;r,2)\}_{k=0}^{\infty}=\cdots=\{p(k;r,r)\}_{k=0}^{\infty}$ holds with probability 1.*

Proof: We first show that if θ_r is ergodic,

$$\mathbf{1}_A = \mathbf{1}_{\theta_1 A} \quad P\text{–a.s.} \quad \text{for all } A\in\mathcal{G}_r. \tag{4.5}$$

For this, observe from the definition of $\mathcal{G}_r$ (Eq. (2.3)) that θ_r being ergodic implies all sets $A\in\mathcal{G}_r$ satisfy $P(A)=1$ or $P(A)=0$. For $A\in\mathcal{G}_r$ satisfying $P(A)=0$, Eq. (4.5) clearly holds. For $A\in\mathcal{G}_r$ satisfying $P(A)=1$, we have

$$\Pr\{\omega:\mathbf{1}_A(\omega)>\mathbf{1}_{\theta_1 A}(\omega)\} = \Pr\{\omega:\mathbf{1}_A(\omega)=1,\mathbf{1}_{\theta_1 A}(\omega)=0\} \tag{4.6}$$
$$\leq \Pr\{\omega:\mathbf{1}_{\theta_1 A}(\omega)=0\}$$
$$= 1-P(\theta_1 A)$$
$$= 0.$$

Similarly, we have $\Pr\{\omega:\mathbf{1}_A(\omega)<\mathbf{1}_{\theta_1 A}(\omega)\}=0$. Therefore, Eq. (4.5) holds for all $A\in\mathcal{G}_r$.

From Lemma 2.1, for all $A\in\mathcal{G}_r$ we have

$$\int_A p(k;r,2)P(d\omega) = \int_A \mathbf{1}(a_1^{(r,2)}=k)P(d\omega) \tag{4.7}$$
$$= \int_A \mathbf{1}(x_2^r=k)P(d\omega) \quad \text{for all } A\in\mathcal{G}_r.$$

Combining $x_2^r(\omega)=x_1^r(\theta_1\omega)$ and Eq. (4.5) yields

$$\int_A p(k;r,2)P(d\omega) = \int_A \mathbf{1}(x_1^r(\theta_1\omega)=k)P(d\omega) = \int_{\theta_1 A}\mathbf{1}(x_1^r=k)P(d\omega) \tag{4.8}$$
$$= \int_A \mathbf{1}(x_1^r=k)P(d\omega) + \left\{\int_{\theta_1 A}-\int_A\right\}\mathbf{1}(x_1^r=k)P(d\omega)$$
$$= \int_A \mathbf{1}(x_1^r=k)P(d\omega) = \int_A p(k;r,1)P(d\omega).$$

By induction, we have

$$\int_A p(k;r,\eta)P(d\omega) = \int_A p(k;r,1)P(d\omega) \quad \text{for all } \eta. \tag{4.9}$$

We define

$$\Theta \equiv \{\omega\in\Omega : p(k;r,\eta)(\omega) > p(k;r,1)(\omega)\}. \tag{4.10}$$

Since $\Theta \in \mathcal{G}_r$, from Eq. (4.9), we get

$$\int_{\Theta} \{p(k;r,\eta) - p(k;r,1)\} P(d\omega) = 0. \tag{4.11}$$

This implies that $P(\Theta) = 0$. Similarly, we have

$$\Pr\{\omega \in \Omega : p(k;r,\eta)(\omega) < p(k;r,1)(\omega)\} = 0. \tag{4.12}$$

This confirms that for all η

$$p(k;r,\eta) = p(k;r,1) \quad P\text{–a.s.} \tag{4.13}$$

proving the desired result. ■

Theorem 4.1 *Suppose that time-shift operator θ_r is ergodic for $(\Omega, \mathcal{F}_x, P)$. Then, for all η*

$$CLR \le \frac{\sum_k [k - \min\{r, K+1\}]^+ p(k;r,\eta)}{\sum_k k p(k;r,\eta)} \quad P\text{–a.s.,} \tag{4.14}$$

where

$$p(k;r,\eta) = \Pr\{a_1^{(r,\eta)} = k\} \quad \textit{for all } \eta, \tag{4.15}$$

and

$$p(k;r,1) = \cdots = p(k;r,r). \tag{4.16}$$

Proof: Corollary 3.1 can be written in the following form.

$$CLR \le \min_{\eta} \left\{ \frac{\sum_k [k - \min\{r, K+1\}]^+ p(k;r,\eta)}{\sum_k k p(k;r,\eta)} + q(r,\eta) \right\}. \tag{4.17}$$

By Lemma 4.2, $\{p(k;r,\eta)\}_{k=0}^{\infty}$ is identical for all η with probability 1 under the given ergodic assumption. Applying Lemma 4.1 to the above, we thus get

$$\begin{aligned} CLR &\le \frac{\sum_k [k - \min\{r, K+1\}]^+ p(k;r,\eta)}{\sum_k k p(k;r,\eta)} + \min_{\eta}\{q(r,\eta)\} \\ &\le \frac{\sum_k [k - \min\{r, K+1\}]^+ p(k;r,\eta)}{\sum_k k p(k;r,\eta)}. \end{aligned} \tag{4.18}$$

Since time-shift operator θ_r is now assumed to be ergodic, all sets $A \in \mathcal{G}_r$ satisfy $P(A) = 1$ or $P(A) = 0$. The conditional expectation on $\mathcal{G}_r$ in Eq. (2.4) is therefore equal to the expectation on the whole set, Ω, that is

$$p(k;r,\eta) = E[\mathbf{1}(a_1^{(r,\eta)} = k) | \mathcal{G}_r] = E[\mathbf{1}(a_1^{(r,\eta)} = k)] = \Pr\{a_1^{(r,\eta)} = k\}. \tag{4.19}$$ ■

Theorem 4.1 implies that Eq. (4.14) holds under the assumption that the stochastic process $\{x_i^r\}_{i=-\infty}^{\infty}$ is ergodic for a time shift of length rL/C. Since the σ-field generated by $\{a_i^{(1)}\}_{i=-\infty}^{\infty}$ includes the σ-field generated by $\{x_i^r\}_{i=-\infty}^{\infty}$, $\{a_i^{(1)}\}_{i=-\infty}^{\infty}$ being ergodic for a time shift of length rL/C is also a sufficient condition under which Eq. (4.14) holds. It is easy to see that Equation (4.14) gives a tighter bound than Eq. (3.4) because $\min\{r, K+1\} \ge \min\{r, (K+2)/2\}$. That is, the cell-arrival process being ergodic makes the cell-loss-ratio upper bound tighter.

5. Numerical Examples

To verify the tightness of the derived upper bounds, we tested several numerical examples. Consider a definition period sequence on the time axis. The 1-st definition period is assumed to start at the 1st slot, that is $\eta = 1$. Cells are assumed to arrive according to two different arrival patterns (Fig. 5). In the first pattern, a batch of cells arrives at the beginning of each definition period. In the second pattern, a batch also arrives in each period, but the arrival alternates between occurring at the beginning and at the end of each period. (In the even-number (odd-number) definition periods, the cells arrive at the end (beginning) of each definition period.) There are two batch sizes: b_1 ($\leq r$) and b_2 ($\geq K+1$). The size changes in accordance with the Markov chain. Let $P = \{p_{ij}\}$ be the transition probability matrix where

$$p_{ij} = \Pr\{s_n = b_j | s_{n-1} = b_i\}, \tag{5.1}$$

and s_n is the batch size in the n-th period. Transition matrix, P, which is given by

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} \tag{5.2}$$

is assumed to be common for both cell-arrival processes. Note that both cell-arrival patterns have a common relative frequency of the number of cells arriving during the definition period; it is expressed as

$$p(k;r,1) = \begin{cases} 1/(1+t) & \text{for } k = b_1 \\ t/(1+t) & \text{for } k = b_2 \\ 0 & \text{otherwise,} \end{cases} \tag{5.3}$$

where $t = \alpha/\beta$.

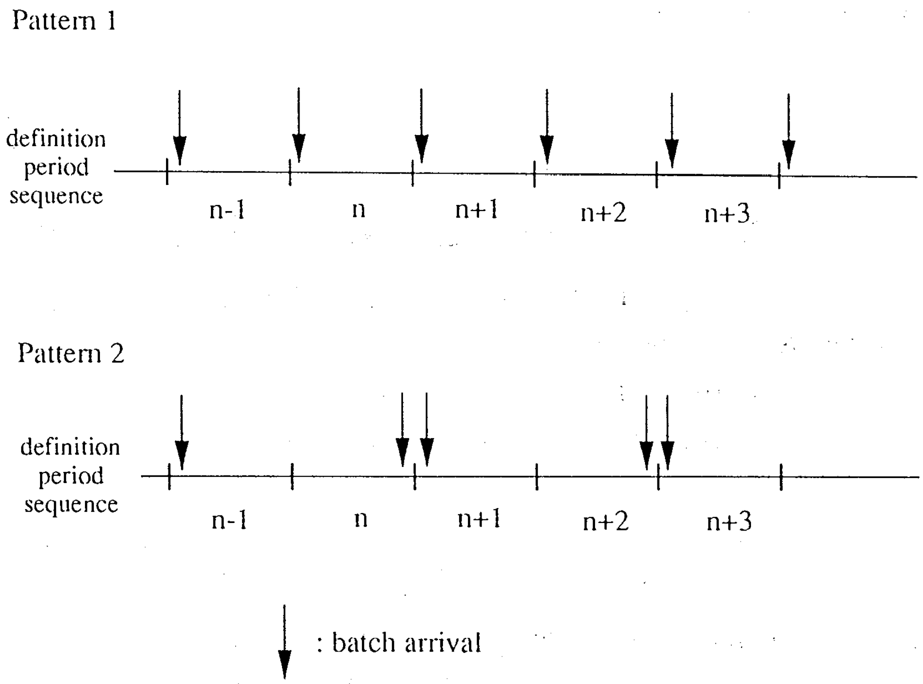

Figure 5. Cell arrival patterns.

The first pattern satisfies the ergodic property discussed in Section 4 ($\{x_i^r\}_{i=-\infty}^{\infty}$ is ergodic for a time shift of length of the definition period). Thus, Eq. (3.4) and Eq. (4.14) are both applicable. The second pattern does not satisfy the ergodic property, so Eq. (3.4) is only applicable to the second pattern.

We can calculate the exact cell-loss ratios when cells arrive according to either of the two cell-arrival patterns at a single-server queue with a finite buffer that serves cells under the first-in-first-serve discipline: if cells arrive according to the first pattern,

(5.4) for $K+1>r$,

$$\begin{aligned}
\boldsymbol{q}_n(0) &= \sum_{l=0}^{r-b_1} \boldsymbol{q}_{n-1}(l)A, \\
\boldsymbol{q}_n(k) &= \boldsymbol{q}_{n-1}(r-b_1+k)A \qquad (0<k\le K-r), \\
\boldsymbol{q}_n(K+1-r) &= \sum_{l=K+1-b_1}^{K} \boldsymbol{q}_{n-1}(l)A + \sum_{l=0}^{K} \boldsymbol{q}_{n-1}(l)B, \\
\boldsymbol{q}_n(k) &= (0,0) \qquad (K+2-r\le k\le K+1),
\end{aligned}$$

and for $K+1\le r$,

$$\begin{aligned}
\boldsymbol{q}_n(0) &= (1/(1+t), t/(1+t)), \\
\boldsymbol{q}_n(k) &= (0,0) \qquad (0<k\le K+1),
\end{aligned}$$

where

(5.5)
$$A=\begin{bmatrix} 1-\alpha & \alpha \\ 0 & 0 \end{bmatrix}, \qquad B=\begin{bmatrix} 0 & 0 \\ \beta & 1-\beta \end{bmatrix},$$
$$\boldsymbol{q}_n(k) = (\Pr\{q_n=k, s_n=b_1\},\ \Pr\{q_n=k, s_n=b_2\}),$$

and q_n denotes the number of cells in the system just before cells arrive in the n-th definition period. If cells arrive according to the second pattern,

(5.6) for $K+1>2r-1$,

$$\begin{aligned}
\boldsymbol{q}_{2n}(0) &= \sum_{l=0}^{2r-2b_1} \boldsymbol{q}_{2(n-1)}(l)A, \\
\boldsymbol{q}_{2n}(k) &= \boldsymbol{q}_{2(n-1)}(2r-2b_1+k)A \qquad (0<k<K+2-2r), \\
\boldsymbol{q}_{2n}(K+2-2r) &= \sum_{l=K+2-2b_1}^{K} \boldsymbol{q}_{2(n-1)}(l)A + \sum_{l=0}^{K} \boldsymbol{q}_{2(n-1)}(l)B, \\
\boldsymbol{q}_{2n}(k) &= (0,0) \qquad (K+3-2r\le k\le K+1),
\end{aligned}$$

and for $K+1\le 2r-1$,

$$\begin{aligned}
\boldsymbol{q}_{2n}(0) &= (1/(1+t), t/(1+t)), \\
\boldsymbol{q}_{2n}(k) &= (0,0) \qquad (0<k\le K+1),
\end{aligned}$$

where

(5.7)
$$A=\begin{bmatrix} (1-\alpha)^2 & \alpha(1-\alpha) \\ 0 & 0 \end{bmatrix}, \qquad B=\begin{bmatrix} \alpha\beta & \alpha(1-\beta) \\ \beta(2-\alpha-\beta) & (1-\beta)^2+\alpha\beta \end{bmatrix},$$
$$\boldsymbol{q}_{2n}(k) = (\Pr\{q_{2n}=k, s_{2n}=b_1\}, \Pr\{q_{2n}=k, s_{2n}=b_{2n}\}).$$

Using the above equations, we numerically calculated the stationary queue-length distribution at a batch-cell arrival epoch:

(5.8)
$$\Pr\{q=k, s=b_i\} = \lim_{n\to\infty} \Pr\{q_n=k, s_n=b_i\}, \text{ and}$$
$$\Pr\{q=k, s=b_i\} = \lim_{n\to\infty} \Pr\{q_{2n}=k, s_{2n}=b_i\}.$$

The cell-loss ratio can be obtained based on the queue-length distribution at a batch-cell arrival epoch.

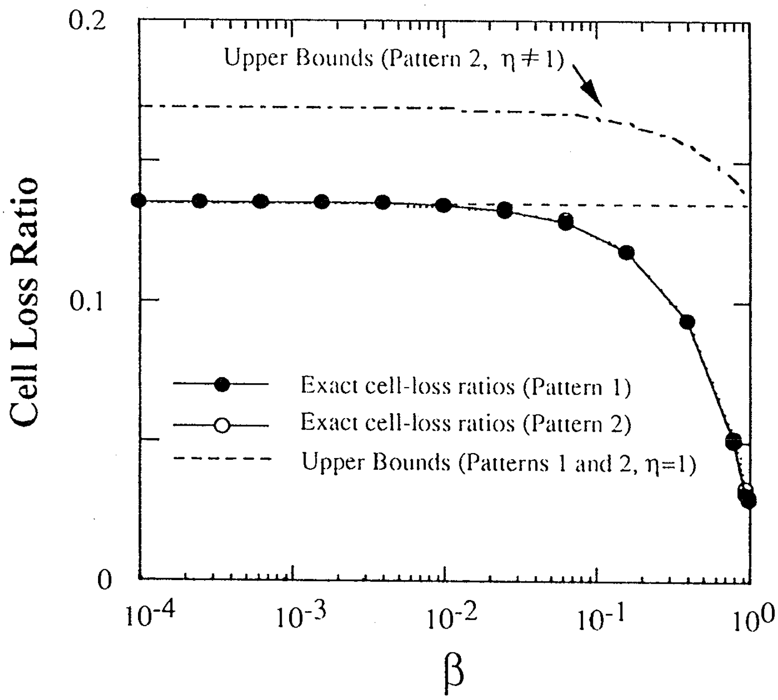

Figure 6a. Comparison between actual cell-loss ratios and upper bounds: $r = 50$.

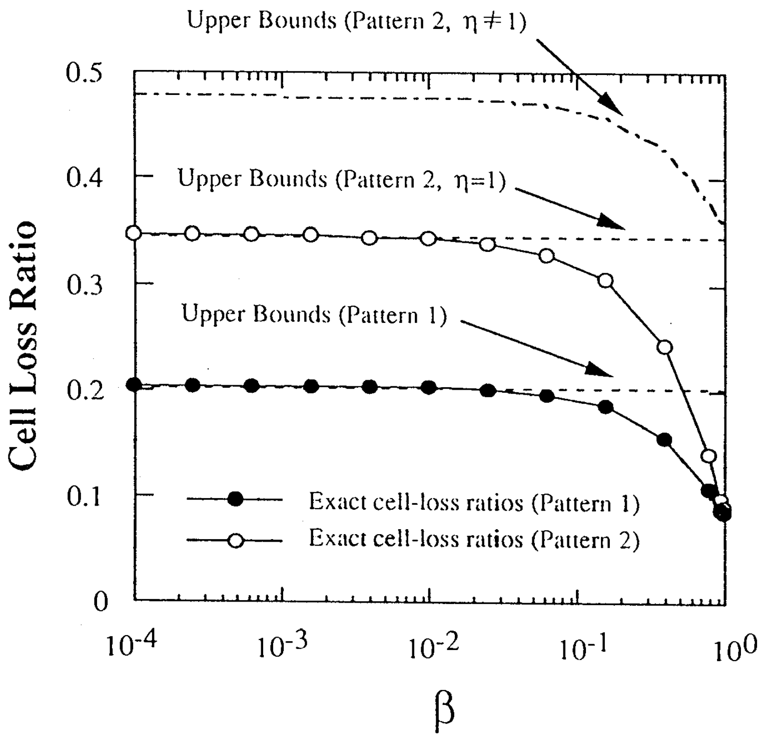

Figure 6b. Comparison between actual cell-loss ratios and upper bounds: $r = 100$.

Figures 6a and 6b compare the exact cell-loss ratios and upper bounds of the derived formulas while changing β from 0 to 1. The value of α was set so as to fix traffic intensity

ρ at 0.05, where ρ is given by

(5.9) $$\rho = \Big(\frac{\beta}{\alpha+\beta}b_1 + \frac{\alpha}{\alpha+\beta}b_2\Big)/r = \frac{b_1+b_2t}{(1+t)r}.$$

Note that since ρ is fixed, t and therefore the relative frequency $p(k;r,1)$ are also fixed while the parameters α and β change. We respectively apply Eqs. (4.14) and (3.4) to the first and second patterns to obtain the upper bounds of the cell-loss ratio. The batch sizes were respectively 2 and 150, and the buffer size was 128. Figure 6a shows the results for $r = 50$, where Eq. (3.4) and Eq. (4.14) give the same upper bound. The exact results approach very close to the upper bounds as $\beta \to 0$. (Letting $\beta \to 0$ corresponds to the case where the same batch size continues for an infinite time.) Figure 6b shows the results for $r = 100$, where Eq. (3.4) and Es. (4.14) give different upper bounds. The exact results for both patterns 1 and 2 approach very close to the upper bounds as $\beta \to 0$. These numerical results confirm that, even though specified by the same relative frequency, cell-arrival processes can have quite different cell-loss ratios and that the derived bounds give the true bounds.

Concerning the second cell arrival pattern, the relative frequency of the number of cells arriving during the definition period depends on η: the relative frequency when $\eta \neq 1$ is given by

(5.10) $$p(k;r,\eta) = \begin{cases} 1/2 & \text{for } k = 0 \\ (1-\beta t)/\{2(1+t)\} & \text{for } k = 2b_1 \\ \beta t/(1+t) & \text{for } k = b_1 + b_2 \\ (1-\beta)t/\{2(1+t)\} & \text{for } k = 2b_2 \\ 0 & \text{otherwise.} \end{cases}$$

(The relative frequency regarding the first cell arrival pattern does not depend on η.) The CLR upper bound therefore also depends on whether $\eta = 1$ or not. The CLR upper bound when $\eta \neq 1$ is also shown in Figs. 6a and 6b. This figure confirms that, although the CLR upper bound depends on the value of η (that is, measurement conditions), this upper bound certainly gives conservative system performance regardless of η.

While this example is rather artificial, it reveals the property of the upper bound very clearly. Other more realistic examples are found in [17].

6. Multiplexed connections and effective bandwidth

In this section, we discuss the "effective bandwidth", which is considered to be one promising solution for ATM resource assignment, based on the derived upper bounds. In particular, we derive the effective bandwidth formula in terms of the relative frequency of the number of cells arriving during a specified interval. For simplicity, we assume that the cell-arrival process is ergodic for a time shift of length rL/C. Under this assumption, the relative frequency of the number of cell arrivals during a fixed interval can be replaced by the probability distribution of the number of cell arrivals, that is,

(6.1) $$p(k;r,\eta) = \Pr\{a_1^{(r,\eta)} = k\}.$$

From Theorem 4.1, the cell-loss ratio is bound as

(6.2) $$CLR \le \frac{\sum_k [k - \min\{r, K+1\}]^+ \Pr\{a_1^{(r)} = k\}}{\sum_{k=0} k\Pr\{a_1^{(r)} = k\}},$$

where we denote $a_1^{(r,\eta)}$ as $a_1^{(r)}$ for notational simplicity because the value of η makes no difference in the following discussion. Consider the case where cells from different connections arrive at the output buffer in an ATM node. These connections are categorized into several classes in terms of their traffic characteristics: the connections categorized into the same class are assumed to have the same traffic characteristics. In such cases,

$$CLR \le \frac{\sum_k [k - \min\{r, K+1\}]^+ \Pr\{\sum_{j=1}^{J}\sum_{i=1}^{n_j} a_1^{(r)}(i,j) = k\}}{\sum_{k=0} k\Pr\{\sum_{j=1}^{J}\sum_{i=1}^{n_j} a_1^{(r)}(i,j) = k\}}. \tag{6.3}$$

Here, $a_1^{(r)}(i,j)$ is the number of cells arriving during the 1-st definition period from the i-th connection of class j, and n_j is the number of connections of class j. Equation (6.3) can be written as

$$CLR \le \frac{\sum_{k=\min\{r,K+1\}+1}^{\infty} \Pr\{\sum_{j=1}^{J}\sum_{i=1}^{n_j} a_1^{(r)}(i,j) \ge k\}}{\sum_{j=1}^{J}\sum_{i=1}^{n_j} E[a_1^{(r)}(i,j)]}. \tag{6.4}$$

Chernoff's bound [11] gives

$$\Pr\{\sum_{j=1}^{J}\sum_{i=1}^{n_j} a_1^{(r)}(i,j) \ge k\} \le e^{-\theta k} \prod_{j=1}^{J}\prod_{i=1}^{n_j} E[\exp(\theta a_1^{(r)}(i,j))] \quad \text{for all } \theta. \tag{6.5}$$

We therefore have

$$\begin{aligned} CLR &\le \frac{\prod_{j=1}^{J}\prod_{i=1}^{n_j} E[\exp(\theta a_1^{(r)}(i,j))] \sum_{k=\min\{r,K+1\}+1}^{\infty} e^{-\theta k}}{\sum_{j=1}^{J}\sum_{i=1}^{n_j} E[a_1^{(r)}(i,j)]} \\ &= \frac{\prod_{j=1}^{J}\prod_{i=1}^{n_j} E[\exp(\theta a_1^{(r)}(i,j))] e^{-\theta(\min\{r,K+1\}+1)}}{\sum_{j=1}^{J}\sum_{i=1}^{n_j} E[a_1^{(r)}(i,j)](1-e^{-\theta})} \\ &= \frac{\prod_{j=1}^{J}\prod_{i=1}^{n_j} \exp(\theta C_j(r,\theta)) e^{-\theta(\min\{r,K+1\}+1)}}{\sum_{j=1}^{J}\sum_{i=1}^{n_j} E[a_1^{(r)}(i,j)](1-e^{-\theta})} \\ &= \frac{\exp\left\{\theta(\sum_{j=1}^{J} n_j C_j(r,\theta) - \min\{r, K+1\})\right\}}{\sum_{j=1}^{J} n_j \hat{A}_j(r)(e^{\theta}-1)} \quad \text{for all } \theta, \end{aligned} \tag{6.6}$$

where

$$(6.7) \qquad C_j(r,\theta) \stackrel{\text{def}}{=} \frac{\log E[\exp(\theta a_1^{(r)}(i,j))]}{\theta}, \quad \text{and} \quad \hat{A}_j(r) \stackrel{\text{def}}{=} E[a_1^{(r)}(i,j)].$$

By taking the logarithm of both sides of the above equation, we get

$$(6.8) \qquad \log(CLR) \le \theta(\sum_{j=1}^{J} n_j C_j(r,\theta) - \min\{r, K+1\}) - \log(e^{\theta}-1) - \log(\sum_{j=1}^{J} n_j \hat{A}_j(r)) \quad \text{for all } \theta.$$

Suppose that θ^* is the value obtained in the infimum of the first term of the right-hand side of Eq. (6.9), that is

$$(6.9) \quad \inf_{\theta} \left[\theta\Big(\sum_{j=1}^{J} n_j C_j(r,\theta) - \min\{r,K+1\}\Big) \right] = \theta^* \sum_{j=1}^{J} (n_j C_j(r,\theta^*) - \min\{r,K+1\}).$$

It can be seen that θ^* is positive when $\sum_{j=1}^{J} \sum_{i=1}^{n_j} E[a_1^{(r)}(i,j)] < \min\{r, K+1\}$ (see Appendix). Obviously, we have

$$(6.10) \quad \log(CLR) \le \theta^*(\sum_{j=1}^{J} n_j C_j(r,\theta^*) - \min\{r,K+1\}) - \log(e^{\theta^*}-1) - \log(\sum_{j=1}^{J} n_j \hat{A}_j(r)).$$

The QOS requirement for cell-loss ratio, CRL_{qos}, is therefore satisfied if

$$(6.11) \qquad \sum_{j=1}^{J} n_j C_j(r,\theta^*) \le \hat{r},$$

where

$$(6.12) \qquad \hat{r} \stackrel{\text{def}}{=} \min\{r,K+1\} + \frac{\log(CLR_{qos}) + \log(e^{\theta^*}-1) + \log(\sum_j n_j \hat{A}_j(r))}{\theta^*}.$$

Equation (6.11) can be interpreted as follows: each connection of class j "effectively" sends $C_j(r,\theta^*)$ cells during r slots – in this sense $C_j(r,\theta^*)$ is called the "effective bandwidth" of the connection. The cell-loss-ratio objective is approximately satisfied if the sum of the effective bandwidths of the multiplexed connections is less than the "effective link capacity" $\hat{r}$ defined by Eq. (6.12) that depends on the cell-loss-ratio objective and the statistics of the connections.

Kelly derived two different relationships regarding the CLR, both of which are similar to Eq. (6.11). The first is based on the effective bandwidth for a bufferless model [9, 3] such that

$$(6.13) \qquad \sum_{j=1}^{J} n_j C_j(1,\theta^{**}) \le 1 + \frac{\log CLR_{qos}}{\theta^{**}},$$

where

$$(6.14) \qquad \inf_{\theta} \left[\theta(\sum_{j=1}^{J} n_j C_j(1,\theta) - 1) \right] = \theta^{**} \sum_{j=1}^{J} (n_j C_j(1,\theta^{**}) - 1).$$

The other is based on the effective bandwidth for the infinite-buffer case (slotted/batch model) [9, 10]:

$$(6.15) \qquad \sum_{j=1}^{J} n_j C_j(r,\zeta) \le r,$$

where

$$\zeta \overset{\text{def}}{=} \frac{\log(CLR_{qos})}{K}. \tag{6.16}$$

Equation (6.13) is very close to Eq. (6.11): the latter is an extended version of the former so as to be applicable still for the case where $r \neq 1$. By contrast, there is a more essential difference between Eq. (6.11) and Eq. (6.15): Eq. (6.15) implies that the connection-acceptable region is linearly constrained in the space of $\boldsymbol{n} = (n_1, n_2, \cdots, n_J)$, and that call acceptance can be judged by simply comparing the sum of the effective bandwidths of multiplexed connections, $\{C_j(r,\zeta)\}_j$, and the link capacity. Because of its conceptual simplicity, the effective bandwidth for the infinite-buffer case has received much attention and many related works have been reported [3, 4, 7, 8, 5]. (The effective bandwidth for the buffer-less model in Eqs. (6.11) and (6.13) is not as helpful as Eq. (6.15) for judging connection acceptance because the effective bandwidth $C_j(r,\theta^*)$ depends on the number of the connections multiplexed in the same queue as well as on the traffic characteristics of the connection.)

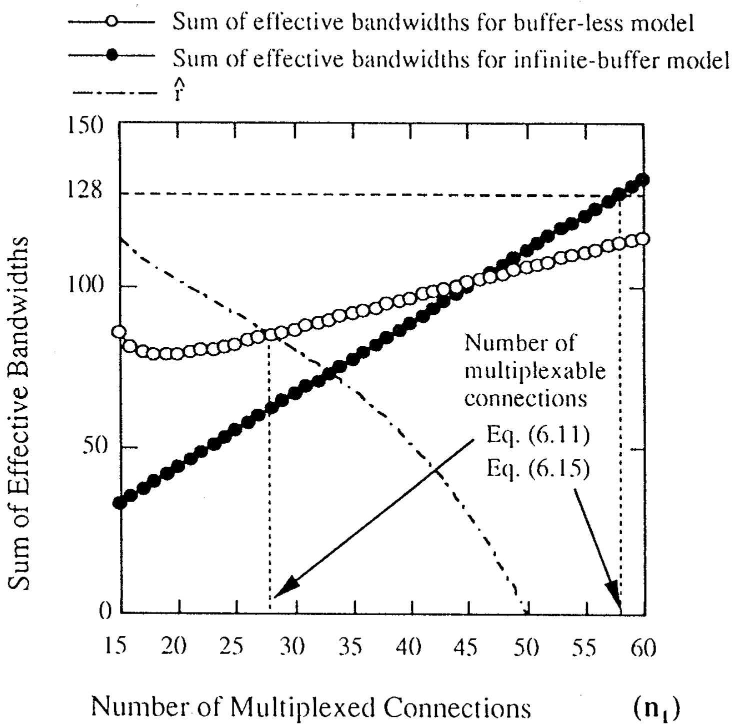

Figure 7. Effective bandwidth and effective link capacity.

Figure 7 quantitatively compares the sum of the effective bandwidths for the infinite-buffer model, $\sum_{j=1}^{J} n_j C_j(r,\zeta)$, and that for the bufferless model, $\sum_{j=1}^{J} n_j C_j(r,\theta^*)$, while changing the number of multiplexed connections. We consider the homogeneous environment (i.e. $J=1$) where virtual channel connections with the peak bit rate = 10 Mb/s and the average bit rate = 2 Mb/s are multiplexed on a 150-Mb/s virtual path. The number of cells arriving during a definition period from the i-th connection is assumed to have the following distribution:

$$\Pr\{a_1^{(r)}(i,1) = k\} = \begin{cases} 1 - A/R & \text{for } k = 0 \\ A/R & \text{for } k = R \\ 0 & \text{otherwise,} \end{cases} \tag{6.17}$$

where A and R respectively denote the average and maximum number of cells arriving during a definition period. That is, $R = \lceil 10r/150 \rceil = \lceil r/15 \rceil$ and $A = 2r/150 = r/75$. Figure 7 shows the case where $r = 128$, $K = 128$, and $CLR_{qos} = 1.0 \times 10^{-4}$. For comparison, $\hat{r}$ defined by Eq. (6.12) is also plotted. The number of multiplexable connections given by Eq. (6.11) is identified in Fig. 7 by the point where the sum of effective bandwidths for the bufferless model is equal to $\hat{r}$, and it is 27. The number of multiplexable connections given by Eq. (6.15) is identified by the point where the sum of effective bandwidths for the infinite-buffer model is equal to r, and it is 57. The number of multiplexable connections is also obtained by using the upper bound formula Eq. (4.14) directly, and results in 33. (The CLR under n multiplexed connections is evaluated by taking n-time convolution of Eq. (6.17) and by substituting the resultant distribution into Eq. (4.14). The number of multiplexable connections is given by the maximum number of multiplexed connections such that the evaluated CLR is less than CLR_{qos}.) Connection acceptance judgement by Eq. (6.15) therefore quite overestimates the number of multiplexable connections in this example. This is because Eq. (6.15) assumes that $(a_1^{(r)}(i,j), a_2^{(r)}(i,j), \cdots, a_n^{(r)}(i,j), \cdots)$ are mutually independent random variables. The cell-loss ratio standard therefore is not always guaranteed by the effective bandwidth approach for an infinite-buffer model like Eq. (6.15), particularly when there is a positive correlation between the number of cells during the consecutive definition periods.

Remark: *the effective bandwidth for the infinite-buffer model can be modified so as to take into account the correlation between the number of cells during the consecutive definition periods* (*see* [3, 4, 7, 8]). *Such an effective bandwidth formula is not, however, obtained only from the distribution of the number of cells during a given interval, and more traffic information about the correlation is required.*

7. Conclusion

We have studied how to evaluate the performance of a queueing system when there is insufficient knowledge of traffic statistics for the case where the relative frequency of the number of cells arriving during a fixed interval is the only available information on traffic characteristics. Insufficient knowledge of the traffic characteristics does not result in an evaluation that is not unique. We focused on the worst-case system performance, that is, the upper bound of the cell-loss ratio when the cell flows are input into a single-server queue with a finite buffer. Since this worst case approach might lead to an overly pesimistic result in some cases, evaluating the performance based on the most-probable behavior probably needs to be studied further.

Appendix

Define

(A.1)
$$f(\theta) \stackrel{\text{def}}{=} \theta\Big(\sum_{j=1}^{J} n_j C_j(r,\theta) - \min\{r, K+1\}\Big)$$
$$= \sum_{j=1}^{J}\sum_{i=1}^{n_j} \log E[\exp(\theta a_1^{(r)}(i,j))] - \theta \min\{r, K+1\}.$$

We first see

(A.2)
$$\frac{df}{d\theta} = \sum_{j=1}^{J}\sum_{i=1}^{n_j} \frac{E[a_1^{(r)}(i,j) e^{\theta a_1^{(r)}(i,j)}]}{E[e^{\theta a_1^{(r)}(i,j)}]} - \min\{r, K+1\},$$

and

$$\text{(A.3)} \qquad \left.\frac{df}{d\theta}\right|_{\theta=0} = \sum_{j=1}^{J}\sum_{i=1}^{n_j} E[a_1^{(r)}(i,j)] - \min\{r, K+1\},$$

$$\text{(A.4)} \qquad \left.\frac{df}{d\theta}\right|_{\theta=\infty} = \sum_{j=1}^{J}\sum_{i=1}^{n_j} \mathrm{Max}[a_1^{(r)}(i,j)] - \min\{r, K+1\},$$

where

$$\text{(A.5)} \qquad \mathrm{Max}[a_1^{(r)}(i,j)] \stackrel{\text{def}}{=} \max_k\{k; \Pr[a_1^{(r)}(i,j) = k] > 0\}.$$

In addition,

$$\text{(A.6)} \quad \frac{d^2 f}{d\theta^2} = \sum_{j=1}^{J}\sum_{i=1}^{n_j} \frac{E[(a_1^{(r)}(i,j))^2 e^{\theta a_1^{(r)}(i,j)}]E[e^{\theta a_1^{(r)}(i,j)}] - E[a_1^{(r)}(i,j)e^{\theta a_1^{(r)}(i,j)}]^2}{E[e^{\theta a_1^{(r)}(i,j)}]^2}$$

$$= \sum_{j=1}^{J}\sum_{i=1}^{n_j} E\left[\left(a_1^{(r)}(i,j) - E\left[a_1^{(r)}(i,j)\frac{e^{\theta a_1^{(r)}(i,j)}}{E[e^{\theta a_1^{(r)}(i,j)}]}\right]\right)^2 \frac{e^{\theta a_1^{(r)}(i,j)}}{E[e^{\theta a_1^{(r)}(i,j)}]}\right] \geq 0.$$

Therefore, $df/d\theta$ is an increasing function for θ. In usual cases, $df/d\theta|_{\theta=\infty}$ is positive because the first term of the right side in Eq. (A.4) is the maximum number of cells arriving during the definition interval. (If it is less than $\min\{r, K+1\}$, clearly $CLR = 0$ and thus it is uninteresting.) Thus if $df/d\theta|_{\theta=0}$ is negative, $df/d\theta = 0$ holds at $\theta^* > 0$. Namely, the function $f(\theta)$ takes its minimum value at $\theta^* > 0$ when

$$\text{(A.7)} \qquad \sum_{j=1}^{J}\sum_{i=1}^{n_j} E[a_1^{(r)}(i,j)] - \min\{r, K+1\} < 0.$$

When $r < K+1$, the above condition is equivalent to a traffic intensity that is less than 1, which is the usual stability condition of the queues.

Acknowledgments

We would like to thank Konosuke Kawashima of NTT Multimedia Networks Laboratories for his valuable advice and discussion. We would also like to thank the referees for their very careful review of this paper. Their insightful comments have led to a greatly improved version of the paper.

References

[1] Baccelli, F. and Bremaud, P., "Elements of Queueing Theory", Springer-Verlag, 1994.
[2] Billingsley, A. A., "Ergodic Theory and Information", John Wiley, New York, 1965.
[3] Botvich, D., and Duffield, N., "Large deviations, the shape of the loss curve, and economies of scale in large multiplexers", *Queueing Systems*, vol. 20, pp. 293–320, 1995.
[4] Chang, C. S., "Stability, queue length, and delay of deterministic and stochastic queueing networks", *IEEE Trans. Auto. Cont.*, vol. 39, no. 5, pp. 913–931, 1994.
[5] Choudhury, G. L., Lucantoni, D. M., and Whitt, W., "On the effectiveness of effective bandwidths for admission control in ATM networks", *ITC 14*, pp. 411–420, 1994.
[6] Doshi, B. T., "Deterministic rule based traffic descriptors for broadband ISDN: worst case behavior and connection acceptance control", *ITC 14*, pp. 591–600, 1994.

[7] Duffield, N. G., "Exponential bounds for queues with Markovian arrivals", *Queueing Systems*, vol. 17, pp. 413–430, 1994.
[8] Elwalid, A. I. and Mitra, D., "Effective bandwidth of general Markovian traffic sources and admission control of high speed networks", *IEEE/ACM Trans. Networking*, vol. 1, no. 3, pp. 329–343, 1993.
[9] Kelly, F. P., "Effective bandwidth at multi-class queues", *Queueing Systems*, vol. 9, pp. 5–16, 1991.
[10] Kelly, F. P., "Tarrifs and effective bandwidths in multiservice networks", *ITC 14*, pp. 401–410, 1994.
[11] Kleinrock, L., "Queueing Systems, Volume I: Theory", John Wiley & Sons, 1975.
[12] Lee, D. C., "Effects of leaky bucket parameters on the average queue delay: worst case analysis", *INFOCOM '94*, pp. 482–489, 1994.
[13] Saito, H., "Call admission control using upper bound of cell loss probability", *IEEE Trans. Commun.*, vol. 40, no. 9, pp. 1512–1521, 1992.
[14] Saito, H., "Dynamic call admission control in ATM networks", *IEEE J. Select. Areas Commun.*, vol. 9, no. 7, pp. 982–989, 1991.
[15] Saito, H., "Toward a future traffic dimensioning method: non-parametric approach for cell loss rate evaluation", *IEICE Trans. Commun.*, vol. J76-B-I, no. 3, pp. 197–208, 1993.
[16] Saito, H., "Teletraffic Technologies in ATM Networks", Artech House, 1994.
[17] Shioda, S. and Saito, H., "A connection admission control guaranteeing a negotiated cell-loss ratio of the cell streams passing through usage parameter control", *IEICE Trans. Commun.* (to appear).
[18] Yamanaka, N., Sato, Y., and Sato, K., "Performance limitation of the leaky bucket algorithm for ATM networks", *IEEE Trans. Commun.*, vol. 43, no. 8, pp. 2298–2230, 1995.

Shigeo Shioda
NTT Multimedia Networks Laboratories
3-9-11 Midoricho, Musashino-shi, Tokyo 180, Japan
e-mail: s-sioda@hashi.tnl.ntt.co.jp

Hiroshi Saito
NTT Multimedia Networks Laboratories
3-9-11 Midoricho, Musashino-shi, Tokyo 180, Japan
e-mail: saito@hashi.tnl.ntt.co.jp