Loss Performance Analysis for Heterogeneous ON–OFF Sources With Application to Connection Admission Control

Guoqiang Mao, *Student Member, IEEE*, and Daryoush Habibi, *Member, IEEE*

***Abstract*—The bufferless fluid flow model (bffm) is often used in the literature for loss performance analysis. In this paper, we propose an efficient and effective means of investigating cell loss using the bffm. We define the cell loss rate function (clrf) and use it to characterize the loss performance of traffic sources in the bffm. Stochastic ordering theory is used to study the clrf. The introduction of the stochastic ordering theory not only simplifies the theoretical analysis but also makes it possible to extend the scope of applications and theoretical analysis presented in this paper. A cell loss upper bound for heterogeneous ON–OFF sources is proposed. The proposed cell loss upper bound is tighter than those previously proposed in the literature. A connection admission control (CAC) scheme using online measurements is designed based on the cell loss upper bound. Extensive simulation is carried out to study the performance of the CAC scheme. Particular attention is paid to the impact of inaccuracies in user-declared traffic parameters on the performance of the CAC scheme. Simulation results indicate that the proposed CAC scheme can ensure QoS guarantee, is robust to inaccuracies in declared traffic parameters, and is capable of achieving high link utilization.**

***Index Terms*—ATM, bufferless fluid flow model, measurement-based CAC, QoS, stochastic ordering theory.**

I. Introduction

TO DATE, many connection admission control (CAC) schemes have been proposed [1]–[8]. These schemes can be classified into two categories: traffic descriptor-based CAC and measurement-based CAC. Traffic descriptor-based CAC uses the *a priori* traffic characterizations provided by sources at connection setup phase to compute whether a new connection in addition to all existing ones can be supported. This approach achieves high network utilization when traffic descriptors used by the CAC scheme are "tight." Measurement-based CAC uses the *a priori* traffic characterizations only for the incoming connection and uses measurements to characterize existing connections. Under the measurement-based CAC scheme, network utilization does not suffer significantly if traffic descriptions are inaccurate. However, because source behavior may be nonstationary, it is difficult for measurement-based CAC to obtain accurate online measurements. Measurement-based CAC can only deliver significant gain in utilization when there is a high degree of statistical multiplexing [1].

Manuscript received October 4, 1999; revised January 29, 2001; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor G. de Veciana.

The authors are with Intelligent Pixel Incorporation, Joondalup, WA 6027, Australia (e-mail: g.mao@ieee.org).

Publisher Item Identifier S 1063-6692(02)01088-9.

Considering the difficulties with both traffic descriptor-based and measurement-based CAC to obtain accurate traffic characteristics, the performance of a CAC scheme should not be measured only by the utilization achieved under ideal circumstances where traffic sources are all tightly characterized. Also, one must consider whether enough accurate traffic characteristics can be obtained from sources and/or network practically, and, the robustness of the CAC scheme against the inaccuracies in those traffic characteristics. In addition to high network utilization, an ideal CAC scheme should satisfy the following requirements [9], [10]:

- *Simplicity*: The scheme must be both economically implementable and fast. The traffic characteristics required by the CAC scheme should be easily and reliably obtained from the traffic sources and/or network.
- *Flexibility*: The scheme must not only be able to satisfy the current needs of network services but also be able to adapt to new services which are likely to evolve.
- *Robustness*: The scheme must be able to handle imperfect assumptions.

Cell loss and cell delay are often adopted as measures of quality of service (QoS). Cell delay can usually be controlled within a desired bound by engineering the buffer size, hence cell loss is used in most papers as the QoS index.

In this paper, based on the bufferless fluid flow model (bffm), we first define the cell loss rate function (clrf) and use it to study the cell loss QoS. To analyze the clrf, we resort to the stochastic ordering theory. The loss performance of heterogeneous ON–OFF sources in bffm is investigated using the clrf and the stochastic ordering theory. A conjecture proposed by Rasmussen *et al.* [11] conceiving cell loss of ON–OFF sources is proved. Furthermore, we propose a tighter cell loss upper bound than that previously proposed in the literature. The upper bound is applied to design a measurement-based CAC scheme. In the design of the CAC scheme, the above guiding principles for CAC schemes are addressed with particular attention given to the impact of inaccuracies in declared traffic parameters.

The rest of the paper is organized as follows. In Section II, the clrf is defined and the stochastic ordering theory is introduced. In Section III, loss performance of heterogeneous ON–OFF sources in bffm is investigated and the cell loss upper bound is developed; the CAC scheme is designed in Section IV. Section V presents some simulation using exponential ON–OFF sources. In Section VI, we analyze the robustness of the CAC scheme against inaccuracies in declared traffic parameters. In

Section VII, video sources are used to test the performance of our CAC scheme, and conclusions are given in Section VIII.

II. Evaluating Cell Loss in the Bufferless Fluid Flow Model—Cell Loss Rate Function

A bffm is often used in the literature to analyze cell loss [7], [12], [13]. Since this approach assumes that there is no buffer at the burst time scale [14], it generates conservative estimates for cell loss. However, the simplicity of bffm enables us to concentrate on the characteristics of traffic sources themselves. In this paper, bffm is employed to analyze the cell loss.

Under bffm, cell loss due to overflow occurs if and only if the sum of the traffic rates of all active connections, denoted by X, exceeds the link capacity C. Let us define a function $F(y)$ as

$$F(y) \triangleq E[(X-y)^+]. \tag{1}$$

We call $F(y)$ the clrf of X. The clrf has many attractive features which facilitate the analysis of cell loss in bffm. For example, $F(C)$ denotes the cell loss rate of a traffic source X on a link with link capacity C. Traffic sources with similar clrf can be regarded as equivalent from the point of view of cell loss analysis. From the definition of clrf, it can be shown that

$$F(y) = F(0) - y, \quad \text{for } y < 0 \tag{2}$$

$$F(0) = E(X). \tag{3}$$

Then, the cell loss ratio (clr) can be calculated as

$$\mathrm{clr} = \frac{F(C)}{F(0)}. \tag{4}$$

An important property of the clrf is given below.

Property 1: If $f(x)$ and $g(x)$ denote the traffic density distribution of independent traffic sources X_1 and X_2, respectively, then the clrf of $X_1 + X_2$ is given by $F * g(x)$, where $F(x)$ is the clrf of X_1, and $*$ denotes convolution.

Proof: Construct a function $h(x)$ such that

$$h(x) = \begin{cases} -x, & x < 0 \\ 0, & x \geq 0. \end{cases}$$

It can be shown that $F(x) = h * f(x)$.

Thus, the clrf of $X_1 + X_2$, denoted by $FG(x)$, is

$$\begin{aligned} FG(x) &= E(X_1 + X_2 - x)^+ \\ &= h * (f * g(x)) \\ &= (h * f) * g(x) \\ &= F * g(x). \end{aligned}$$

■

To analyze the clrf, we turn to the stochastic ordering theory. Given two random variables X and Y, we say that X is smaller than Y with respect to the *increasing convex ordering*, written as $X <_{\mathrm{icx}} Y$, if for the distribution function of X and Y, denoted by F_d and G_d, the following condition:

$$\int_{-\infty}^{+\infty} \phi(x)\, dF_d(x) \leq \int_{-\infty}^{+\infty} \phi(x)\, dG_d(x) \tag{5}$$

holds for all increasing convex function ϕ, for which the integral exists [15]. The following property of the increasing convex ordering enables us to use it to analyze the clrf.

Lemma 1: $X <_{\mathrm{icx}} Y$ if and only if $E[(X-x)^+] \leq E[(Y-x)^+]$ for any $x \in R$.

See [15, Sect. 1.3, Theorem A] for a proof. According to Lemma 1, $X <_{\mathrm{icx}} Y$ if and only if the clrf of X is smaller than or equal to the clrf of Y for any real value.

An important *special case* of the increasing convex ordering is when X and Y have the same mean value, i.e., $E(X) = E(Y)$. In this case, we say that X is smaller than Y with respect to the *convex ordering*, written as $X <_{\mathrm{cx}} Y$, because the characterizing inequality (5) holds for all convex functions [16, Corollary 8.5.2]. $X <_{\mathrm{cx}} Y$ implies that not only the cell loss of X is less than or equal to that of Y, but also, from (3) and (4), the clr of X is less than or equal to that of Y. We refer to [15]–[17] for the properties of the increasing convex ordering.

III. Heterogeneous ON–OFF Sources in the Bufferless Fluid Flow Model

For simplicity, let us consider ON–OFF traffic sources. According to the probability density distribution of ON and OFF periods, ON–OFF source models can be further classified into exponential ON–OFF source model, periodic ON–OFF source model, Pareto ON–OFF source model, etc. They are widely used for loss performance analysis [7], [11], [12], [18], [19].

An ON–OFF source generates cells at a peak cell rate (pcr) denoted by pcr in active periods. In idle periods, no cells are generated. Let mcr denote the mean cell rate (mcr) of an ON–OFF source. The activity parameter p of an ON-OFF source is defined as the ratio of mcr to pcr: $p \triangleq \mathrm{mcr}/\mathrm{pcr}$. The probability that an ON–OFF source is active or idle is given by p or $1-p$, respectively.

Assume there are n independent ON–OFF sources $X_1, \ldots, X_n$ on the link, where $X_{i(i=1,\ldots,n)}$ has peak cell rate pcr_i, mean cell rate mcr_i and activity parameter $p_i = \mathrm{mcr}_i/\mathrm{pcr}_i$. Denote the probability mass function (pmf) of $X_{i(i=1,\ldots,n)}$ by $f_{1,(\mathrm{pcr}_i)}^{(p_i)}(x)$. The pmf of $\sum_{i=1}^{n} X_i$ can be expressed as

$$f_{n,(\mathrm{pcr}_1,\ldots,\mathrm{pcr}_n)}^{(p_1,\ldots,p_n)}(x) = f_{1,(\mathrm{pcr}_1)}^{(p_1)} * \cdots * f_{1,(\mathrm{pcr}_n)}^{(p_n)}(x). \tag{6}$$

In this paper, we use subscript n, subscript (pcr), and superscript (p) to denote the number of ON–OFF sources, peak cell rates of sources, and their activity parameters, respectively, when it is necessary to emphasize the dependence of a function on these parameters.

Since an ON–OFF source or the multiplexing of ON–OFF sources, denoted by X, is a discrete random variable, the definition in (1) can be simplified as

$$F(y) \triangleq E[(X-y)^+] \triangleq \sum_x (x-y)^+ f(x) \tag{7}$$

where $f(x)$ is the probability mass function of the discrete random variable X.

We shall now introduce an important property of ON–OFF sources.

Proposition 1: Assume X is an arbitrary traffic source such that $E(X) = \mathrm{mcr}$ and $\|X\|_\infty = \mathrm{pcr}$, where $\|X\|_\infty = \inf\{x :$

$\Pr\{X > x\} = 0\}$. Let Y represent an ON–OFF source with mean cell rate mcr and peak cell rate pcr. Then $X <_{\mathrm{cv}} Y$.

It has been shown in [17, App. B] that $X <_{\mathrm{icx}} Y$. Then, noting that $E(X) = E(Y)$, Proposition 1 is easily proved. This proposition implies that, among traffic sources with the same mcr and pcr, ON–OFF source constitutes the worst-case in cell loss analysis.

We shall now introduce an important theorem about heterogeneous Bernoulli sources. Bernoulli sources are ON–OFF sources with the same peak cell rate.

Theorem 1: Let $X_1, \ldots, X_n$ be n independent heterogeneous Bernoulli sources with the same peak cell rate and activity parameters $p_1, \ldots, p_n$, respectively. The activity parameters are subject to $p_1 + \cdots + p_n = P$. Let $Y_1, \ldots, Y_n$ be n independent homogeneous Bernoulli sources, where $Y_{i(i=1,\ldots,n)}$ has the same peak cell rate as X_i and an activity parameter $p = P/n$. Then $\sum_{i=1}^n X_i <_{\mathrm{cx}} \sum_{i=1}^n Y_i$.

See Appendix A for a proof.

Theorem 1 states that homogeneous Bernoulli sources generate more cell loss than that of heterogeneous Bernoulli sources. Theorem 1 was first proposed as a conjecture by Rasmussen *et al.*[11]. Based on Theorem 1, they proposed that the cell loss of n heterogeneous ON–OFF sources, whose maximum peak cell rate is pcr, is upper bounded by that of n homogeneous ON–OFF sources with peak cell rate pcr, where the sum of mean cell rates remains the same. In addition to [11], Theorem 1 is also used for loss performance analysis in many other literature [7], [20], [21].

In real networks, many traffic sources of the same type have the same pcr. However, because of their specific application circumstances, they may have different mean cell rates. Theorem 1 is very useful for analyzing this kind of traffic source.

We shall now introduce another theorem about ON–OFF sources.

Theorem 2: Let X and Y be two independent ON-OFF sources with the same mean cell rate denoted by mcr. X and Y have peak cell rates pcr_X and pcr_Y respectively. If $\mathrm{pcr}_X \le \mathrm{pcr}_Y$, then $X <_{\mathrm{cx}} Y$.

Proof: Since $E(X) = E(Y)$, it suffices to show that the clrf of Y is greater than or equal to that of X. This can be easily shown by computing the clrf of X and Y and comparing their values. ■

In [12], a result to the same effect of Theorem 2 is proved using another approach.

Theorem 2 is used by Lee *et al.* [12] in their CAC scheme. Based on it, they design a CAC scheme capable of real-time estimation of cell loss of the multiplexing of heterogeneous ON–OFF sources. Theorem 2 will be used for loss performance analysis and CAC scheme design in this paper.

On the basis of the theorems, lemmas, and properties shown previously, we shall now introduce an important theorem about independent heterogeneous ON–OFF sources.

Theorem 3: Let $X_1, \ldots, X_n$ be n independent heterogeneous ON–OFF sources with peak cell rates $\mathrm{pcr}_1, \ldots, \mathrm{pcr}_n$ and mean cell rates $\mathrm{mcr}_1, \ldots, \mathrm{mcr}_n$, respectively. Let $Y_1, \ldots, Y_m$ represent m independent homogeneous ON–OFF sources with peak cell rate pcr and mean cell rate mcr, where $\mathrm{pcr} \ge \max\{\mathrm{pcr}_1, \ldots, \mathrm{pcr}_n\}, m = \lceil ((\sum_{i=1}^n \mathrm{pcr}_i)/\mathrm{pcr}) \rceil$ and $\mathrm{mcr} = ((\sum_{i=1}^n \mathrm{mcr}_i)/m)$. Then

$$\sum_{i=1}^n X_i <_{\mathrm{cv}} \sum_{i=1}^m Y_i \tag{8}$$

For a proof, see Appendix B.

This theorem provides a cell loss upper bound for heterogeneous ON–OFF sources. It states that the cell loss of heterogeneous ON–OFF sources is less than or equal to that of the corresponding homogeneous ON–OFF sources given in the theorem. The sum of mean cell rates of the homogeneous sources remains the same as that of the heterogeneous sources, and the sum of peak cell rates of the homogeneous sources is also substantially the same as that of the heterogeneous sources. It is not difficult, using the clrf and Theorem 1, to prove that our upper bound is tighter than the upper bound proposed by Rassussen *et al.* [11].

IV. CAC Scheme

The proposed upper bound can be applied to either traffic descriptor-based CAC or measurement-based CAC. Here we choose to design a measurement-based CAC.

Measurement-based CAC has attracted a lot of interest in recent years [1], [4], [7], [8], [22]–[26]. Shiomoto *et al.* [23] use a lowpass filter to obtain the instantaneous virtual path (VP) utilization from crude measurements. A residual bandwidth is derived from the maximum of the observed instantaneous VP utilization. If the bandwidth requirement of the new connection is smaller than the residual bandwidth, the new connection is admitted, otherwise it is rejected. Gibbens *et al.* [7] use a decision-theoretic approach for call admission control to explicitly incorporate call-level dynamics into the model. In their work, call admission decisions are based on whether the current measured load is less than a precomputed threshold. In a study by Dziong *et al.*[24], a Kalman filter is used to obtain an optimal estimate of mcr and variance. An aggregate equivalent bandwidth is then obtained from the mean and variance estimates. A spare bandwidth is set in the aggregate equivalent bandwidth to account for the estimation errors. Grossglauser *et al.* [22] study a robust measurement-based admission control with emphasis on the impact of estimation errors, measurement memory, call-level dynamics and separation of time scales. Their work [22], [25] also identifies a *critical time-scale* $\widetilde{T_h}$ such that aggregate traffic fluctuation slower than $\widetilde{T_h}$ can be tracked by the admission controller and compensated for by connection admissions and departures. Fluctuations faster than $\widetilde{T_h}$ have to be absorbed by reserving spare bandwidth on the link. Using Gaussian and heavy traffic approximations, the critical time scale is shown to scale as $T_h/\sqrt{n}$, where T_h is the average flow duration and n is the size of the link in terms of number of flows it can carry.

In this section, based on the loss performance analysis in the previous sections, we shall design a CAC scheme using parameters from both traffic descriptors and measurements. The principles introduced in Section I are used to govern the CAC scheme design. Robustness, flexibility, and simplicity become major concerns in our CAC design and later simulation validation.

Let us select a traffic rate unit u such that u is greater than or equal to the maximum pcr of all traffic sources on the link. From now on in this paper, all traffic rates, as well as link capacity, are normalized with respect to u unless otherwise specified. Without loss of generality, we assume that link capacity C is an integer.

Suppose there are n independent heterogeneous ON–OFF sources, denoted by $X_1, \ldots, X_n$, on the link. The declared pcr of ON–OFF source $X_{i(i=1,\ldots,n)}$ is pcr_i. We keep a list of the declared peak cell rates of all connections on the link and denote the sum of the peak cell rates by PCR, i.e., $\mathrm{PCR} = \sum_{i=1}^{n} \mathrm{pcr}_i$. Realizing that it is difficult for traffic sources to tightly characterize their mean cell rates, we obtain the sum of mean cell rates from online measurements. This is of course the mcr of the link. We denote the measured mcr of the link by r.

From Theorem 3

$$\sum_{i=1}^{n} X_i <_{\mathrm{cv}} \sum_{i=1}^{m} Y_i$$

where

$$m = \lceil \mathrm{PCR} \rceil \tag{9}$$

and $Y_{i(i=1,\ldots,m)}$ is an independent ON–OFF source with peak cell rate 1 and activity parameter p. The choice of p should make the mcr of $\sum_{i=1}^{m} Y_i$ and the mcr of $\sum_{i=1}^{n} X_i$ equal. The method of estimating p from measurements is described later. The pmf of $\sum_{i=1}^{m} Y_i$ is given by the following binomial distribution:

$$f(x) = \begin{cases} \binom{m}{k} p^k (1-p)^{m-k}, & x = k \\ 0, & \text{else.} \end{cases} \tag{10}$$

According to the definition of clrf for discrete random variable given in (7), the clrf of $\sum_{i=1}^{m} Y_i$ can be calculated as

$$F(k) = \begin{cases} m \times p, & k = 0 \\ F(k-1) - 1 + \sum_{i=0}^{k-1} f(i), & k \geq 1 \end{cases} \tag{11}$$

The clr of $\sum_{i=1}^{m} Y_i$ is estimated using (4). It can be shown that clr of $\sum_{i=1}^{m} Y_i$ is greater than or equal to that of $\sum_{i=1}^{n} X_i$.

A. Estimation of the Activity Parameter p

We shall now describe the estimation of p. Since we derive the mcr of $\sum_{i=1}^{n} X_i$ from online measurements, p can be directly estimated as follows:

$$\hat{p} = \frac{r}{m}. \tag{12}$$

Increasing the measurement window size T_m will increase the accuracy of the measured mcr r and the accuracy of estimation of p. Yamada *et al.* [27] introduce a method for finding the measurement window size. Estimation of p using (12) is simple; however, for accurate estimation it requires a large measurement window size. Here we use another approach.

In an ATM network, traffic can only arrive in integer multiples of an ATM cell. Therefore, we first choose a sampling period T_s such that the impact of the granularity of the traffic rate measurements taken over T_s can be ignored. In our analysis, T_s is chosen to be 100 cell unit time. One cell unit time is the time required to transmit an ATM cell on the link. Denote the mean and the variance of the traffic rate sample r_1 measured over T_s by S_T and σ_T^2, and the mean and the variance of the traffic rate sample r_K measured over a sampling period of $K \times T_s$ by S_K and σ_K^2, respectively. Assuming that the aggregate traffic is stationary, then it can be shown that S_K, σ_K, S_T, and σ_T are related by $S_K = S_T$ and

$$\sigma_K^2 = \sigma_T^2 \left[\frac{1}{K} + \frac{1}{K^2} \sum_{i=1}^{K} (K-i)\rho_i \right]$$

where ρ_i is the autocorrelation between traffic rate samples taken over $[0, T_s]$ and over $[(i-1)T_s, i \times T_s]$. σ_K decreases with the increase of K. In the above analysis, the impact of call level dynamics was ignored, i.e., we assume no call is admitted into the network or departs from the network during the measurement window. The impact of call level dynamics is discussed later in this subsection.

According to the Central Limit Theorem, when the number of connections is large, the distribution of the aggregate traffic can be well approximated by Gaussian distribution. Therefore, the aggregate traffic is assumed to be Gaussian. If the following equation is chosen as an estimate of the mcr:

$$\hat{r} = r_K + \varepsilon \times \sigma_K$$

where ε is a constant, in order to satisfy the estimation objective

$$P(\hat{r} \geq S_K) \geq 0.95 \tag{13}$$

we have to choose $\varepsilon = 1.65$. If the aggregate traffic is not Gaussian, a larger ε can be obtained using the Chebyshev's inequality. In the rest of the paper, we only consider the case when the aggregate traffic is Gaussian traffic.

An estimate of σ_K can either be obtained directly from on-line traffic measurements, or can be obtained from on-line estimation of σ_T and autocorrelation function ρ. However, on-line estimation of these parameters is not an easy task, so we take another approach.

The variance of $\sum_{i=1}^{m} Y_i$ is given by $m(S_T/m)(1 - (S_T/m))$. Moreover, using the properties of the convex ordering, it can be easily shown that, if $\sum_{i=1}^{n} X_i <_{\mathrm{cv}} \sum_{i=1}^{m} Y_i$, then the variance of $\sum_{i=1}^{m} Y_i$ is greater than or equal to that of $\sum_{i=1}^{n} X_i$. Therefore, $m(S_T/m)(1-(S_T/m))$ is the maximum variance of the aggregate traffic $\sum_{i=1}^{n} X_i$. So, instead of measuring the variance directly, σ_T is estimated as

$$\widehat{\sigma_T} = \sqrt{m \frac{r_K}{m} \left(1 - \frac{r_K}{m}\right)}. \tag{14}$$

When a sufficiently large K is chosen such that traffic fluctuations with time scale larger than $K \times T_s$ are small, the estimated σ_T is larger than its true value despite possible fluctuations in r_K. An estimate of σ_K^2 is then obtained as

$$\widehat{\sigma_K}^2 = \frac{\widehat{\sigma_T}^2}{K^\delta} \tag{15}$$

where δ is a constant between 0 and 1. Parameter δ can be obtained by inspecting the variance-time plot obtained from traffic measurements [28], [29]. In contrast to [28] and [29] which study the self-similarity in network traffic, our interest is mainly in the variance-time curve in a relatively small time region from $10T_s$ to $1000T_s$. In this region, even for short-range dependent traffic, a δ much smaller than 1 may be observed. In our simulation shown later, δ is a very stable value. This is because the traffic mix in the simulation is almost time-invariant. As a result, the autocorrelation in network traffic does not change significantly. However, our analysis on traffic data from real ATM networks[1] shows that the value of δ will change slowly with time within a day. In that case, the smallest value of observed δ should be used in (15). For simulation using exponential ON–OFF sources shown later, δ is chosen to be 0.35; for simulation using Motion-JPEG encoded video sources, δ is chosen to be 0.4. The advantage of this method is that the value of δ can be obtained from off-line traffic analysis, therefore, online estimation of the second-order statistics is avoided. Here we would like to comment that this method will not give an accurate estimate of σ_K. However, generally the estimated value of σ_K is larger than its true value, which will satisfy the estimation objective in (13).

Then an estimate of p can be obtained: $\hat{p} = (\hat{r}/m)$.

In summary, in the measurement scheme, parameter T_s is chosen to be 100 cell unit time and parameters K and δ can be obtained empirically from off-line traffic analysis. The measurement window size is chosen to be $T_m = K \times T_s$, and an estimate of p is obtained as

$$\hat{p} = \frac{r}{m} + \alpha\sqrt{\frac{\frac{r}{m}\left(1 - \frac{r}{m}\right)}{m}} \tag{16}$$

where

$$\alpha = 1.65 \times K^{-\frac{\delta}{2}} \tag{17}$$

and r is the mean traffic rate measured over T_m.

This approach first appeared in [7]. The introduction of the safety margin α enables us to greatly reduce the required measurement window size while maintaining the robustness of the estimation. Moreover, the safety margin introduces additional benefits. Loss performance analysis presented in this paper gives a cell loss upper bound of the aggregate traffic based on bffm. The measurement scheme shown above also gives a robust estimate of p. Therefore, the above measurement scheme will give a tight QoS guarantee. However, for a network with large buffers, clr will decrease due to large buffer size. The proposed CAC scheme is conservative for a network with large buffers. In this case, safety margin α can be chosen to be smaller than that given in (17), or even zero, to achieve a higher link utilization. Our simulation shows that, for a fixed buffer size, controlling the safety margin α can control the clr, as well as utilization. Selecting a safety margin α to adapt to a network with large buffers is subject to further study.

Therefore, the introduction of α brings some flexibility into the CAC scheme which enables us to efficiently utilize network resources. Thus, we use (16) to estimate p in our CAC scheme instead of (12).

[1]These traffic traces have been collected by Waikatp Applied Network Dynamics group at the University of Auckland since November 1999. The timestamp of measurement data in the trace has an accuracy of 1 μs. More details about the traces can be found at http://moat.nlanr.net/Traces/Kiwitraces.

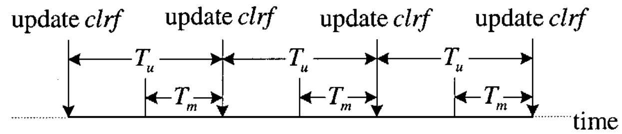

Fig. 1. Relationship between update interval and measurement interval.

The clrf of $\sum_{i=1}^{m} Y_i$ is updated using (10) and (11) every T_u seconds. More specifically, the clrf is updated at the end of each measurement period. Fig. 1 shows the relationship between the update period T_u and measurement period T_m. The update period T_u is chosen empirically. We suggest choosing a T_u in the range $2T_m \sim 10T_m$, depending on the network state. When call level dynamics are fast, i.e., connections enter and leave the network very often, T_u should be chosen close to $2T_m$. On the other hand, for a network where call level dynamics are slow, T_u should be chosen close to $10T_m$.

Call level dynamics affect the measurements. If a new connection is admitted within a given measurement window, the new connection possibly generates traffic only during part of the measurement window or does not generate traffic at all. Therefore, the new connection will make the measured mcr of the link smaller than its actual value, which will affect the robustness of the CAC scheme with regard to QoS guarantees. To solve this problem, when the measurement window size T_m is much smaller than the connection setup time, we block the admission of new connections during the measurement window. The admission will be delayed till the end of the measurement window. In the worst case, this will introduce a delay of T_m to the connection setup time. This method is used in our simulation. Alternatively, if the delay caused by T_m to the connection setup time becomes a concern, one can add the sum of declared mean cell rates, if available in the traffic descriptor, or the sum of declared peak cell rates of the connections admitted in the measurement window, divided by m, to the estimated p in (16). In this case, we do not need to block the admission of new connections in the measurement window.

It is also possible that during the measurement window some existing connections are released, thus affecting the measured mcr. Departing connections contribute to the measured mcr of the link. However, PCR is the sum of peak cell rates of connections on the link at the instant of updating clrf, not including peak cell rates of the departing connections. So if there are some connections that are released during the measurement interval they will make the estimated p larger, which in turn makes the CAC scheme more conservative, assuring that QoS guarantees are not affected. Another alternative is to update the clrf only when no existing connections depart during the measurement interval. We do not adopt this approach in our CAC scheme because it may result in the clrf not being updated for a long time, thus affecting the performance of the CAC scheme.

In our CAC scheme, we do not update the clrf for departing connections. The changes in traffic parameters due to departing connections are caught up by updating the clrf at the end of each measurement period.

B. Cell Loss Ratio Estimation

We shall now present the method to estimate the clr when a new connection request arrives. Suppose there are M connections, denoted by $Z_1, \ldots, Z_M$, admitted into the network since the last measurement interval. Denote by $X_1, \ldots, X_n$ the connections on the link at the instant when the last measurement interval finished, and denote by $F(y)$ the clrf of the corresponding upper bound of $X_1, \ldots, X_n, \sum_{i=1}^{m} Y_i$, which is computed using (9)–(11) and (16). When a new connection request arrives, denoted by Z_{M+1}, if $\mathrm{PCR} + \sum_{i=1}^{M+1} \mathrm{pcr}_{Z_i} \le C$, where $\mathrm{pcr}_{Z_{i(i=1,\ldots,M+1)}}$ is the declared pcr of Z_i, the new connection can be admitted directly and no cell loss will occur. Otherwise, the following method is used to estimate the clr if the new connection is accepted and to determine whether the new connection should be accepted.

Denote the updated clrf after the admission of Z_M by $F_M(y)$. The procedure of updating the clrf is described later. Denote the pcr and mcr of the new connection request Z_{M+1} by $\mathrm{pcr}_{Z_{M+1}}$ and $\mathrm{mcr}_{Z_{M+1}}$. Z_{M+1} is smaller than, with respect to the convex ordering, an ON–OFF source λ with a peak cell rate 1 and mean cell rate $\mathrm{mcr}_{Z_{M+1}}$. Therefore, the cell loss ratio of the aggregate traffic, if the new connection is admitted, is less than

$$\begin{aligned}\mathrm{clr} &= \frac{g * F_M(C)}{g * F_M(0)}\\ &= \frac{\left(1 - \mathrm{mcr}_{Z_{M+1}}\right) F_M(C) + \mathrm{mcr}_{Z_{M+1}} F_M(C-1)}{F_M(0) + \mathrm{mcr}_{Z_{M+1}}}\end{aligned} \tag{18}$$

where g denotes the pmf of λ. If clr is less than the clr objective then the connection is admitted; otherwise the connection is rejected. If the connection is admitted, clrf will be updated:

$$F_{M+1}(k) = \begin{cases} F_M(0) + \mathrm{mcr}_{Z_{M+1}}, & k = 0 \\ \left(1 - \mathrm{mcr}_{Z_{M+1}}\right) F_M(k) & \\ +\mathrm{mcr}_{Z_{M+1}} F_M(k-1), & k \ge 1. \end{cases} \tag{19}$$

For the special case of $M = 0$, the clrf $F_0(y)$ is actually $F(y)$, the clrf of $\sum_{i=1}^{m} Y_i$.

Alternatively, one can also take $\hat{r} + \sum_{i=1}^{M+1} \mathrm{mcr}_{Z_i}$ as an estimate of the sum of mean cell rates of all connections on the link if the new connection request Z_{M+1} is admitted; and calculate the clrf $F_{M+1}(y)$ using (9)–(11) and (16). Accordingly, in (9), pcr now means the sum of peak cell rates of all connection on the link if connection request Z_{M+1} is admitted.

Updating the clrf using (19) is computationally much more efficient. However, since (19) actually takes the peak cell rate of $Z_1, \ldots, Z_M, Z_{M+1}$ as 1, it will generate more conservative results. In the later simulation, (19) is chosen to update the clrf.

Noting that in the estimation of clr only the computation of $F(y)$and $f(x)$ from 0 to C is needed, we do not need to calculate all values of $F(y)$ and $f(x)$ from 0 to $\lceil \mathrm{PCR} \rceil$.

In the above method, the traffic sources are required to specify their peak cell rates and mean cell rates. Using the loss performance analysis presented in this paper, other methods can also be developed which only need the traffic sources to specify their peak cell rates, or need less computation [30].

C. Clr of Individual Connection

In the previous sections, a CAC scheme was developed which is able to guarantee the clr of the aggregate traffic. By choosing the clr objective of the aggregate traffic in the CAC scheme appropriately, the clr requirement of individual connection can also be satisfied.

Assume there are n connections $X_1, \ldots, X_n$ on the link. $X_1, \ldots, X_n$ may have any distribution. The clr of connection $X_{k(k=1,\ldots,n)}$ can be evaluated as

$$\mathrm{clr}_{X_k} = \frac{E\left[\left(\sum_{i=1}^{n} X_i - C\right)^+ \frac{X_k}{\sum_{i=1}^{n} X_i}\right]}{E(X_k)}.$$

Let $Y_1, \ldots Y_n$ be n independent ON–OFF sources where ON–OFF source $Y_{k(k=1,\ldots,n)}$ has the same pcr and mcr as X_k. It is shown in [13, Theorem 1] that the clr of connection X_k is upper bounded by

$$\begin{aligned}\mathrm{clr}_{X_k} \le \mathrm{clr}_k &= \frac{E\left[\left(\sum_{i=1, i\ne k}^{n} X_i + \mathrm{pcr}_k - C\right)^+\right]}{C}\\ &= \frac{F(C - \mathrm{pcr}_k)}{C}\end{aligned}$$

where $F(y)$ is the clrf of $\sum_{i=1, i\ne k}^{n} Y_i$. The clr of the aggregate traffic $\sum_{i=1}^{n} Y_i$, denoted by clr_Y, is related to clr_k by

$$\begin{aligned}\mathrm{clr}_Y &= \frac{p_k \times F(C - \mathrm{pcr}_k) + (1 - p_k) \times F(C)}{E\left(\sum_{i=1}^{n} X_i\right)}\\ &= \frac{p_k \times C \times \mathrm{clr}_k + (1 - p_k) \times F(C)}{E\left(\sum_{i=1}^{n} X_i\right)}.\end{aligned}$$

Denote the clr objective of X_k by $\mathrm{clr}_{k,\mathrm{obj}}$. If

$$\mathrm{clr}_Y \le \frac{p_k \times C \times \mathrm{clr}_{k,\mathrm{obj}} + (1 - p_k) \times F(C)}{E\left(\sum_{i=1}^{n} X_i\right)} \tag{20}$$

is satisfied, the QoS of connection X_k can be guaranteed. We can further remove the term $(1 - p_k) \times F(C)$ from (20), i.e., if

$$\mathrm{clr}_Y \le \frac{p_k \times C}{E\left(\sum_{i=1}^{n} X_i\right)} \times \mathrm{clr}_{k,\mathrm{obj}} \tag{21}$$

is satisfied, then the QoS requirement of connection X_k can be guaranteed. Equation (21) gives more conservative clr objective for the aggregate traffic than that given by (20), however (21) is much easier to implement practically. Since $F(C) < C \times \mathrm{clr}_{k,\mathrm{obj}}$, when p_k is not a small value, the clr objective for the aggregate traffic given by (21) is close to that given by (20).

Therefore, if the clr objective of the aggregate traffic is chosen to be $((p_k \times C)/(E(\sum_{i=1}^{n} X_i))) \times \mathrm{clr}_{k,\mathrm{obj}}$, the proposed CAC scheme is able to guarantee the clr requirement of connection X_k. Also, if the clr objective of the aggregate traffic is chosen to be

$$\inf_{1 \le k \le n} \left(\frac{p_k \times C}{E\left(\sum_{i=1}^{n} X_i\right)} \times \mathrm{clr}_{k,\mathrm{obj}}\right) \tag{22}$$

the clr requirements of all connections can be satisfied.

V. Simulation Study

In this section, we study the performance of the proposed CAC scheme using simulation. The aim of the simulation is to evaluate the performance of our CAC scheme with respect to link utilization and its effectiveness in terms of its ability to guarantee the QoS constrains required by the connections.

The simulation is carried out using OPNET. The following parameters are used in the simulation unless otherwise specified: switching speed of the ATM switch is set to be infinity, hence every incoming cell is placed immediately in the output buffer; the output buffer size is set to be 20 cells to absorb cell level congestion [31], [14]. The link utilization and clr are observed in the simulation. Link utilization is calculated as the ratio of the instantaneous link traffic rate to the link capacity; clr is calculated as the ratio of the total observed cell loss to the total cells offered to the link in a moving window with size T_c. More specifically, clr at time t is the ratio of the number of cell loss occurred in the interval $(t - T_c, t]$ to the total number of cells offered to the link for transmission in the same interval, where T_c equals 500s. In each scenario, there are several types of traffic sources multiplexed onto the link. The connection arrival process of each type of traffic is a Poisson process with a mean of λ calls per second. The connection holding time for all traffic types is exponentially distributed with a mean of 100 s.

A. Simulation Model

In this section we shall use an exponential ON–OFF source model in the simulation. The duration of the ON and OFF periods are independent and exponentially distributed with means β and γ, respectively. During ON periods, cells are generated at peak cell rate. During off periods no cells are generated.

Furthermore, the following parameters are used for the simulation: the link capacity is set to be 10 Mb/s, and the measurement window size is chosen to be 0.08 s. Clrf update period is chosen to be 0.2 s. The safety margin α is chosen to be 1.0. In the simulation, three types of traffic sources are multiplexed on the link. The traffic rate unit u is set to be 100 kb/s, which is the maximum pcr of the three traffic types.

Two scenarios were simulated. In the first scenario, referred to as the saturation scenario, the connection arrival rate is chosen to be very high. The high call arrival rate means that the system is continuously receiving new connection requests. Thus, the CAC scheme is expected to achieve the maximum utilization in the saturation scenario. This scenario is used to establish the performance of our CAC scheme with regard to QoS guarantees, because if calls are offered at a very high rate, the rate at which calls are admitted in error becomes very large too [7]. In the second scenario, referred to as the moderate scenario, the call arrival rate of each traffic type is carefully chosen to make the call blocking ratio fall between 0–0.03. The call blocking ratio is defined as the ratio of the number of calls rejected to the total number of call arrivals. Since real networks are not likely to operate in the saturation scenario, the utilization achieved in the moderate scenario can better represent the utilization that can be achieved by our CAC scheme.

TABLE I
PARAMETERS OF THE THREE TRAFFIC TYPES IN THE SATURATION SCENARIO

	$\lambda\,(s^{-1})$	$pcr\,(kb/s)$	$\beta\,(s)$	$\gamma\,(s)$
type 1	10	100	0.424	3.816
type 2	50	50	0.424	1.696
type 3	100	10	0.424	0.424

B. Saturation Scenario

The parameters of the three types of traffic for the saturation scenario are listed in Table I.

The mean on time of the three traffic types shown in Table I implies that traffic type 1 has a mean burst length of 100 cells, traffic type 2 has a mean burst length of 50 cells, and traffic type 3 has a mean burst length of 20 cells. The mean burst length is several times larger than the buffer size. This is used as a trial to establish the performance of the CAC scheme using online measurement.

The clr objective of the aggregate traffic is set to be 10^{-4}. According to our analysis in Section IV-C, approximating the term $(C/(E(\sum_{i=1}^{n} X_i)))$ by 1, it can be shown that by choosing such an aggregate traffic clr objective, the clr of traffic type 1 can be controlled below 10^{-3}; the clr of traffic type 2 can be controlled below 5×10^{-4} and the clr of traffic type 3 can be controlled below 2×10^{-4}. These parameters are examined in the simulation to validate the proposed CAC scheme as well as our loss performance analysis on individual connections in Section IV-C. It should be noticed that these parameters are only used in the simulation to test the performance of the proposed CAC scheme. In real applications, the clr objective of the aggregate traffic should be chosen according to the clr requirements of individual connections using (22) in order to satisfy the QoS requirements of each connection on the link.

The simulation is run for 10 000 s. An average utilization of 0.76 is achieved by our CAC scheme. Figs. 2 and 3 show the clr of each traffic type as well as the clr of the aggregate traffic. As shown in Figs. 2 and 3, both the clr of the aggregate traffic and the clr of each traffic type are controlled within the desired bound. Fig. 4 shows the number of connections of each traffic type on the link. Fig. 5 shows the admissible region for the three types of connections as well as the number of connections actually admitted by the CAC scheme. The number of admitted connections is close to the boundary of the admissible region but within the admissible region. Therefore, the CAC scheme is robust with regard to QoS guarantees, and is capable of achieving a high link utilization. This is also verified by the observed clr shown in Fig. 2 and 3.

According to Gibbens *et al.* [32] and Guerin *et al.* [33], the effective bandwidth of an ON–OFF source is given by

$$\hat{c} = \frac{\eta\beta(1-p) \times \mathrm{pcr} - B}{2\eta B(1-p)} + \frac{\sqrt{[\eta\beta(1-p)\mathrm{pcr} - B]^2 + 4B\eta\beta p(1-p)\mathrm{pcr}}}{2\eta B(1-p)}$$

where $\eta = \ln(1/\mathrm{clr})$, clr is the cell loss ratio objective, B is the buffer size, and β is the mean on time. Define the statistical multiplexing gain to be the ratio of the bandwidth required to

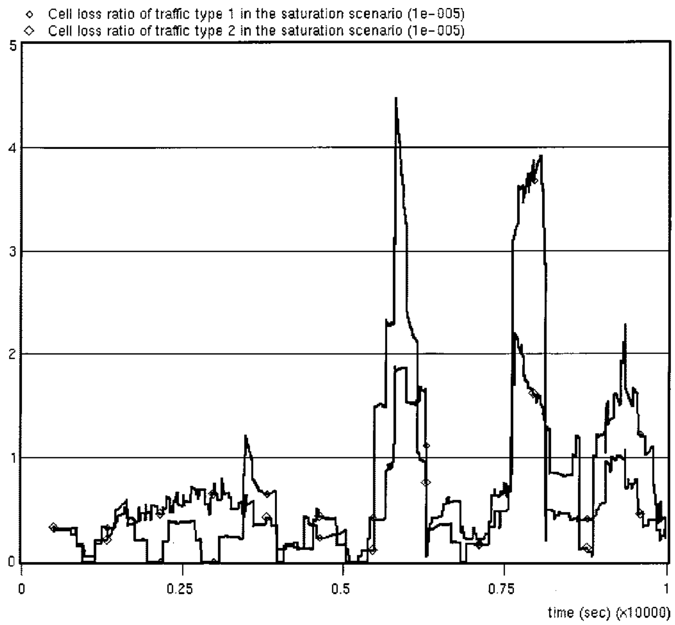

Fig. 2. Cell loss ratio of traffic type 1 and traffic type 2 in the saturation scenario.

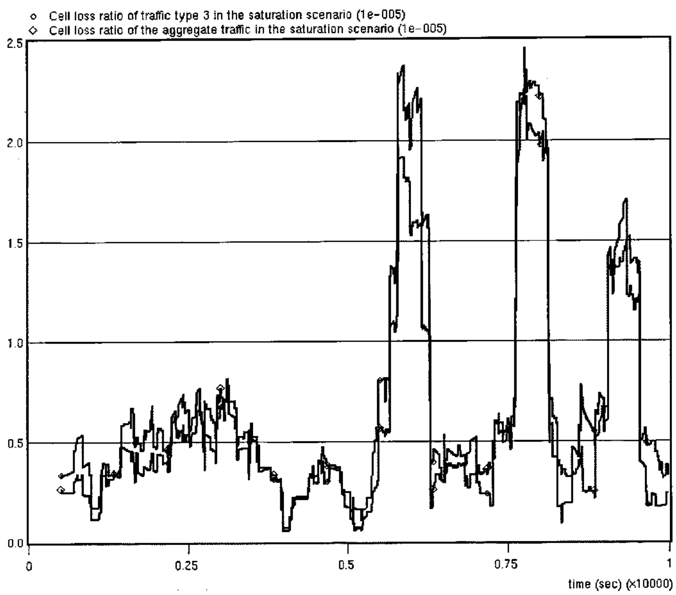

Fig. 3. Cell loss ratio of traffic type 3 and the aggregate traffic in the saturation scenario.

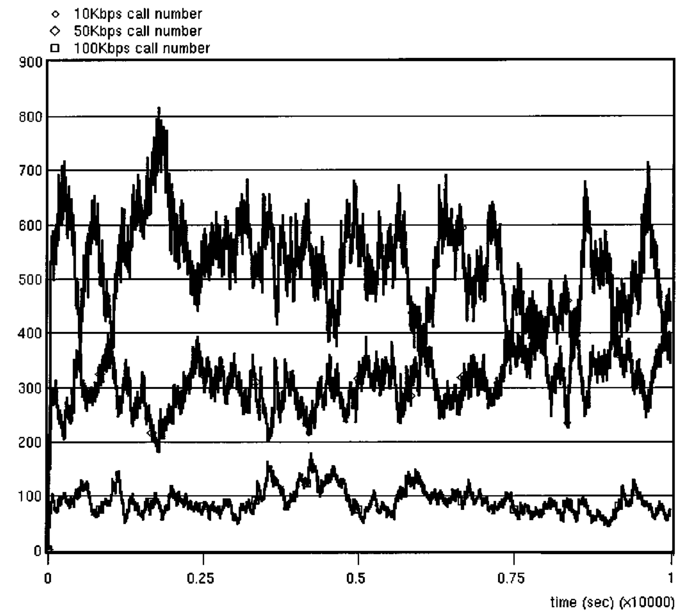

Fig. 4. The number of connections of the three traffic types on the link in the saturation scenario.

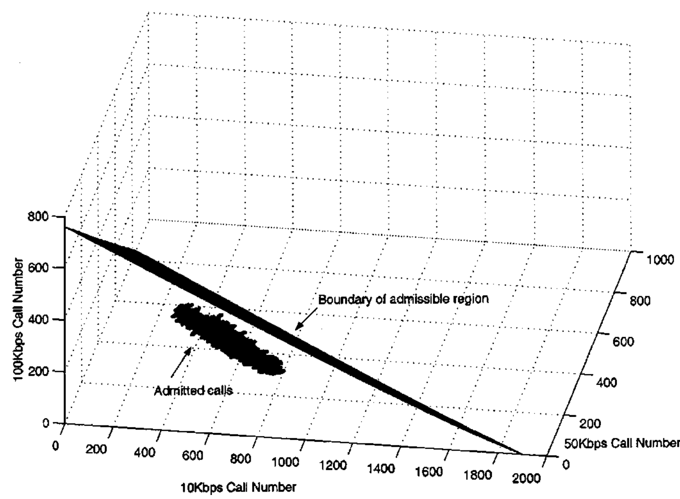

Fig. 5. Comparison between admissible region and actually admitted calls.

support the same number of connections in the simulation using the effective bandwidth approach to the link capacity, which indicates the bandwidth required by our CAC scheme. An average statistical multiplexing gain of 2.7 is achieved in the saturation scenario.

C. Moderate Scenario

In this scenario, we study the case where connections of each type arrive at a moderate rate. The mean values of connection arrival rates of type 1, type 2, and type 3 connections are 0.202, 0.756, and 1.512 call/s, respectively. All other parameters are chosen to be the same as those in the saturation scenario. Again, the simulation was run for 10 000 s. Simulation results show that both the clr of the aggregate traffic and the clr of each traffic type are controlled within the desired bound in the moderate scenario. For ease of comparison, only the clr of the aggregate traffic is presented. Figs. 6 and 7 show the clr of the aggregate traffic and the number of each traffic type on the link.

An average statistical multiplexing gain of 2.1 is achieved in the moderate scenario, with an average utilization of 0.68. Compared with the saturation scenario, link utilization decreased by 0.08. CLR observed in the moderate scenario also decreased. This is a natural consequence of the decreased utilization. In our measurement-based CAC scheme, the estimated mean traffic rate is greater than its true value in most cases. While such an arrangement can ensure robust QoS guarantees, as a penalty, it will inevitably result in false rejections of connections. In the saturation scenario, the false rejections are compensated by high connection arrival rates, thus the utilization is unaffected. However, in the moderate scenario, the decrease in utilization is in-

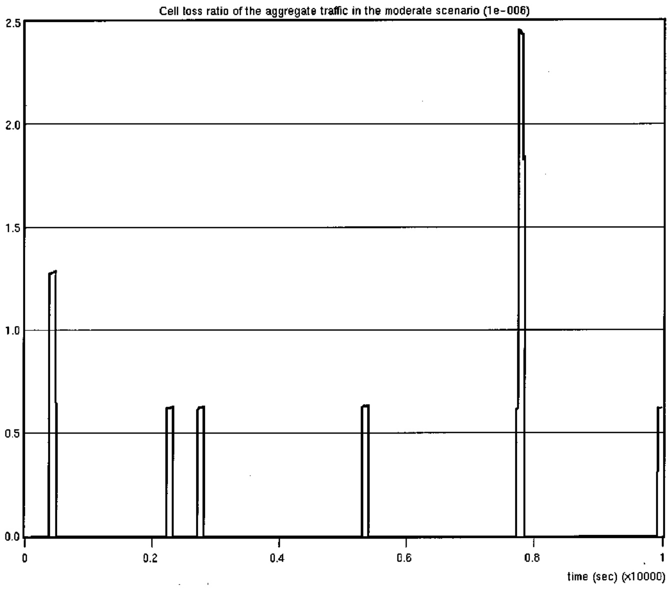

Fig. 6. Cell loss ratio of the aggregate traffic in the moderate scenario.

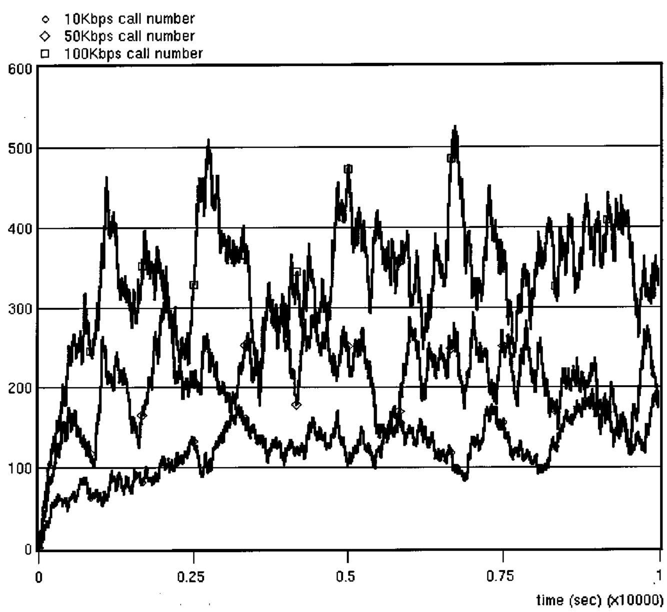

Fig. 7. The number of the three traffic types on the link observed in the moderate scenario.

evitable. The safety margin α can be chosen to be smaller to achieve better utilization, but as a penalty, the degree of certainty of the CAC scheme with regard to QoS guarantees will be decreased in the presence of high connection arrival rate.

VI. Robustness of the CAC Scheme

In Section V, traffic parameters specified by the traffic sources are tight and accurate. However, this is impossible in real networks. Therefore, it is essential for a CAC scheme to be robust against inaccuracies in the declared traffic parameters. In this section, we study the performance of our CAC scheme when declared traffic parameters are not tight.

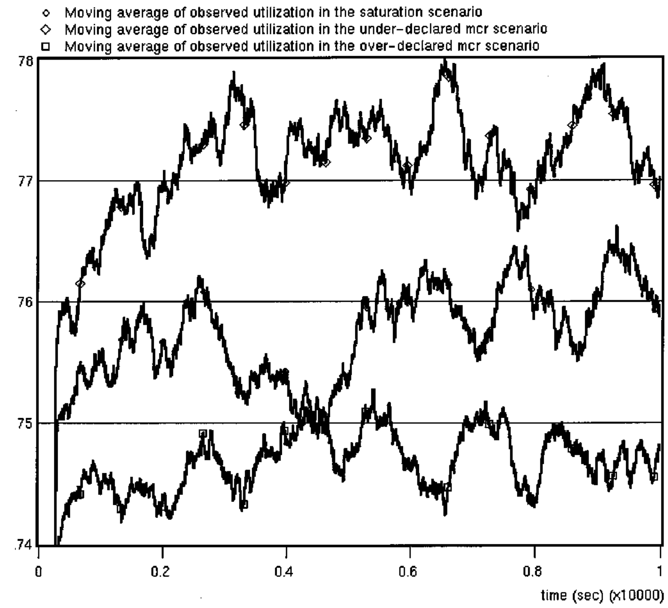

Fig. 8. Impact of inaccuracies in declared mcr on utilization.

Our scheme requires that traffic sources declare their mean cell rates and peak cell rates. The impact of inaccuracies in the declared mcr and pcr is studied separately in this section. The simulation parameters are the same as those in the saturation scenario unless otherwise specified. For comparison purposes, utilization shown in this section is the moving average of observed utilization, average window size is 500 s; clr at time t is the ratio of the total number of cell losses occurred in the interval $[0, t]$ to the total number of cells offered to the link in the same interval. For ease of comparison, only the clr of the aggregate traffic is considered in this section.

A. Impact of Inaccuracies in the Declared Mean Cell Rate

In this subsection we study the impact of inaccuracies in the declared mean cell rates on the performance of the CAC scheme. Two more scenarios are considered in this subsection. In one scenario, referred to as the over-declared mcr scenario, the declared mean cell rates of all three traffic types in the traffic contract are set to be 1.5 times their actual values. In the other scenario, referred to as the under-declared mcr scenario, the declared mean cell rates are set to be 0.5 times their actual values. The performance of the CAC scheme in the saturation scenario, the under-declared mcr scenario and the over-declared mcr scenario is compared.

Fig. 8 shows the moving average of the observed utilization in the three scenarios. Fig. 9 shows the clr of the aggregate traffic. Comparing the utilization and the clr in the three scenarios, it is observed that varying the declared mcr in such a large range only results in very small variation in utilization, i.e., less than 0.01, and slight variation in clr. These results indicate that our CAC scheme is robust against inaccuracies in the declared mcr. This is not unexpected, because inaccuracies in the declared mcr only have a localized effect on the performance of the CAC scheme, that is, its impact is limited to one clrf update interval following which the declared mcr will be replaced by the measured value.

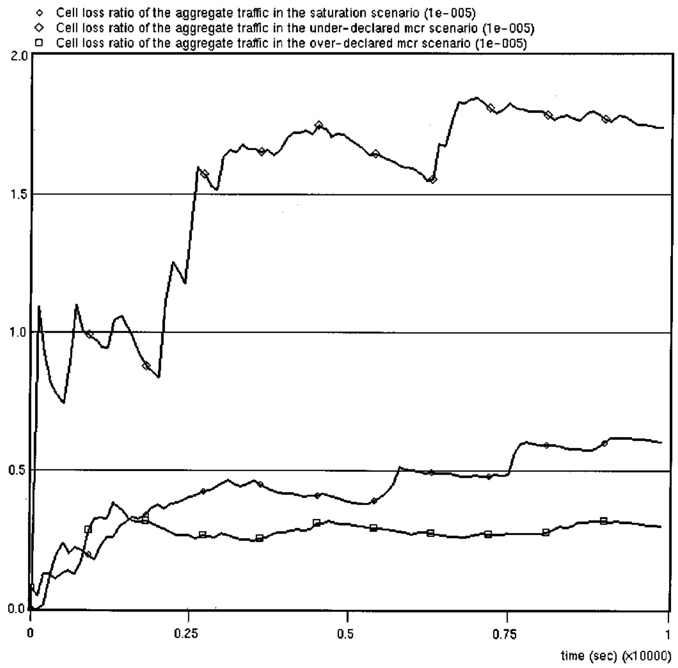

Fig. 9. Impact of inaccuracies in declared mcr on cell loss ratio.

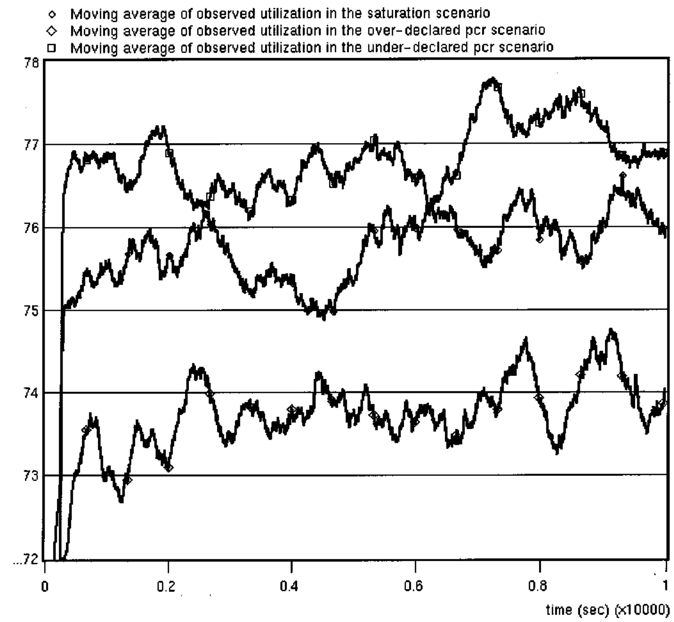

Fig. 10. Impact of inaccuracies in declared pcr on utilization.

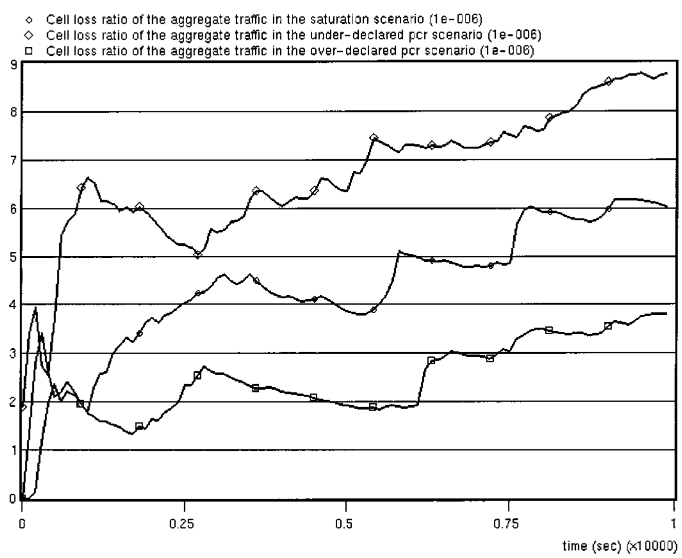

Fig. 11. Impact of inaccuracies in declared pcr on cell loss ratio.

B. Impact of Inaccuracies in the Declared Peak Cell Rate

In this subsection, we study the impact of inaccuracies in the declared pcr on the performance of the CAC scheme. In contrast to the mcr for which inaccuracies in the declared value only have localized effects, the inaccuracies in the declared pcr have long term effects on the performance of our CAC scheme and will persist for the duration of the connection. Hence, it is very important that the CAC scheme be robust against inaccuracies in the declared pcr.

To assess this problem, we consider two new simulation scenarios. In the first scenario, referred to as the over-declared pcr scenario, the declared peak cell rates of all traffic types in the traffic contract are set to be two times their actual values, and, in the second scenario, referred to as the under-declared pcr scenario, the declared peak cell rates are set to be 0.75 times the actual values. The performance of the CAC scheme in the saturation scenario, the under-declared pcr scenario, and the over-declared pcr scenario is compared.

Fig. 10 shows the moving average of the observed utilization for the three scenarios. Fig. 11 shows the clr of the aggregate traffic. In the under-declared pcr scenario, an increase of 0.01 in utilization and a slight increase in clr were observed. In the over-declared pcr scenario, utilization decreases by 0.02 and clr decreases to half of its value in the saturation scenario. These results are very encouraging, i.e., although the declared pcr is varied over a large range, the link utilization and clr do not suffer significantly. These results indicate that the CAC scheme is able to achieve a good performance even when the declared pcr is very inaccurate. We offer the following explanation for the robustness of the CAC scheme against inaccuracies in the declared pcr: over-specifying the pcr will usually decrease link utilization, but it also causes the estimated activity parameter p in our CAC scheme to drop by the same percentage, which leads to an increase in utilization. The combination of these two effects makes link utilization less sensitive to the changes in the declared pcr. When traffic sources under-specify their peak cell rates the reverse process occurs. It is this mechanism which makes our CAC scheme robust against inaccuracies in the declared pcr.

Simulation results and analysis presented in this section show that the proposed CAC scheme is robust against inaccuracies in declared traffic parameters. This feature is very attractive for real applications as it lessens the burden on traffic sources to tightly characterize their traffic parameters. In the simulation, we assume that all traffic sources underspecify or overspecify their traffic parameters. This assumption is only used to establish the performance of the CAC scheme against inaccuracies in the declared traffic parameters. In real networks, the inaccuracies in the declared traffic parameters will possibly offset each other and hence make the sum of the declared traffic parameters

TABLE II
TRAFFIC RATE OF THE M-JPEG ENCODED MOVIES (bytes/frame)

Type	Name	peak rate	mean rate
1	Sleepless in Seattle	16617	9477.6
2	Crocodile Dundee	19439	10772.9
3	Home Alone, II	22009	11382.8
4	Jurassic Park	23883	11363.0
5	Rookie of the Year	27877	12434.9
6	Speed	29385	12374.4
7	Hot Shots, Part Duex	29933	12766.1
8	Beauty and the Beast	30367	12661.5

more accurate. Therefore, in real applications, the impact of inaccuracies in declared traffic parameters will possibly be even less severe than that presented in this section.

VII. APPLICATION OF THE CAC SCHEME TO REAL TRAFFIC SOURCES

In the previous section, we studied the performance of our CAC scheme using the exponential ON–OFF source model. In this section, we will further study the performance of our CAC scheme using variable bit rate video sources. Eight Motion-JPEG (M-JPEG) encoded movies are used in the simulation. Accordingly, there are eight traffic types. Connections of each type have an exponentially distributed duration with a mean of 100 s. Connection arrival process of each traffic type is a Poisson process with a mean of 1 call/s. When a connection of a given type is admitted, it starts reading the corresponding M-JPEG encoded movie file from the beginning and generates traffic according to the movie file. The statistics of the M-JPEG encoded movies are shown in Table II. The frame rate of the M-JPEG encoded movies is 30 frames/s. Details about the M-JPEG encoded movies can be found in [34]. The traffic rate unit u is chosen to be 7.288 Mb/s, which is the maximum pcr of the video sources. An OC3 link is used in the simulation. The measurement window size is chosen to be 0.10 s. Clrf update period is chosen to be 0.2 s. Safety margin α is chosen to be 0.5. The cell loss ratio objective is chosen to be 10^{-4}. The simulation is run for 6 000 s. An average link utilization of 0.65 is achieved in this scenario while the cell loss ratio is controlled within the desired region. The achieved utilization is smaller that that achieved using the exponential ON–OFF sources. There are mainly two factors affecting the utilization. First, the peak cell rates of the traffic sources are close to 10% of the link rate. As a rule of thumb, when the peak cell rates of the traffic sources are close to 10% of the link rate, the statistical multiplexing gain which can be achieved is small. Second, the ON–OFF source model at the basis of our CAC scheme is the worst case among all traffic sources with the same peak and mean. In this scenario, each video source is modeled by an ON–OFF source, which is actually the worst case of the real source. This will also result in lower utilization. It is worth noting that the aggregate traffic in this simulation scenario presents significant self-similar behavior with a Hurst parameter of 0.7 [29]. Therefore, it is possible that the proposed measurement-based CAC scheme can be applied to self-similar traffic. Further study is required to clarify this problem.

VIII. CONCLUSION

In this paper, we proposed the cell loss rate function for studying cell loss in the bufferless fluid flow model. The clrf enabled us to decompose the complex analysis of the multiplexing of traffic sources into simpler analysis of individual sources. Furthermore, stochastic ordering theory was used to analyze the clrf. The introduction of the stochastic ordering theory not only simplified the theoretical analysis but also made it possible to extend the application of theoretical analysis presented in this paper to a broader area. Based on the clrf and the stochastic ordering theory, loss performance of heterogeneous ON–OFF sources was studied. A cell loss upper bound for heterogeneous ON–OFF sources was proposed which may simplify both theoretical analysis and computation. We believe this theoretical analysis forms a good basis for studying ON–OFF sources in the bffm. The methodology presented in this paper also constitute a good starting point for studying other traffic source models in the bffm.

The proposed upper bound can be applied to both parameter-based CAC and measurement-based CAC. In this paper, we used the upper bound to design a measurement-based CAC. Many practical factors were considered in the design of the CAC scheme. Extensive simulation studies were carried out which were indicative of good performance of the CAC scheme. Our CAC scheme proves to be robust against inaccuracies in declared traffic parameters and capable of achieving high link utilization. The proposed CAC scheme has many attractive features which make it suitable for implementation in real ATM networks.

The loss performance analysis presented in this paper is a steady-state analysis, thus the proposed upper bound constitutes the worst case in terms of average loss performance. Theoretically speaking, the average loss performance constraints may not be meaningful, if the aggregate traffic exhibits long range dependence. Specifically, if the aggregate traffic exhibits long-range dependence although average performance may be deemed to be fine, there may be rare periods of time in which performance is consistently poor.

However, there is considerable debate about the impact of long-range dependent traffic on bandwidth allocation and network performance [35]–[38], especially the impact of long-range dependent traffic on the performance of a measurement-based CAC scheme [22], [25], because the time scale of interest in real applications is limited. For measurement-based CAC, Grossglauser *et al.* [22], [25] identify a critical time-scale $\widetilde{T_h}$ such that aggregate traffic fluctuations slower than $\widetilde{T_h}$ can be tracked by the admission controller and compensated for by connection admissions and departures. Following their analysis, only those traffic fluctuations with time scale greater than measurement window size T_m but less than $\widetilde{T_h}$ will threaten the performance of our CAC scheme. Our CAC scheme is based on an upper bound, i.e., it will over-allocate bandwidth. Thus, we consider that if traffic fluctuations with time scale greater than measurement window size T_m but less than $\widetilde{T_h}$ are negligible, or their bandwidth requirement can be satisfied by over-allocated bandwidth, then long-range dependent traffic will not affect the performance of the proposed CAC scheme. Otherwise, the proposed CAC scheme will fail to provide QoS guarantees in the presence of long-range dependent traffic. The impact of the long-range dependent traffic on the performance of our measurement-based CAC is in fact also a problem for almost all measurement-based CAC. Further studies are required to clarify this problem.

APPENDIX A

In this appendix, we shall prove Theorem 1. First we introduce a lemma which will be used in the proof.

Lemma 2: Let X_1, X_2 be two independent ON–OFF sources with peak cell rates $\alpha_1 \times \mathrm{pcr}, (1-\alpha_1) \times \mathrm{pcr}$, and activity parameters p_1, p_2, respectively, where $1 > \alpha_1 \geq 0.5$. Let Y_1 and Y_2 be two independent ON–OFF sources with peak cell rates $\alpha_2 \times \mathrm{pcr}, (1-\alpha_2) \times \mathrm{pcr}$, and activity parameters q_1, q_2, respectively, where $1 > \alpha_2 \geq 0.5$. Moreover, $E(X_1 + X_2) = E(Y_1 + Y_2)$ and $\alpha_1 \leq \alpha_2$. Then $X_1 + X_2 <_{\mathrm{cx}} Y_1 + Y_2$ if and only if $p_1 p_2 \leq q_1 q_2$ and $p_1 + p_2 - p_1 p_2 \geq q_1 + q_2 - q_1 q_2$.

See [30,Ch. 3] for a proof.

Let us now start the proof of Theorem 1.

Proof: It is easy to show that

$$E\left(\sum_{i=1}^{n} X_i\right) = E\left(\sum_{i=1}^{n} Y_i\right).$$

Then, from Lemma 1, it suffices to show that the clrf of $\sum_{i=1}^{n} Y_i$ is greater than or equal to that of $\sum_{i=1}^{n} X_i$ for any real value. The proof is by induction on n.

1) First let us consider the case where $n = 2$. We must show that the clrf of $X_1 + X_2$ is less than or equal to that of $Y_1 + Y_2$, where X_1, X_2 are two independent heterogeneous Bernoulli source with activity parameters p_1, p_2, and the same peak cell rate pcr respectively; and Y_1, Y_2 are two homogeneous Bernoulli sources with activity parameter $p = (p_1 + p_2)/2$ and peak cell rate pcr. It can be shown that

$$p^2 = \frac{(p_1 + p_2)^2}{4} \geq p_1 p_2.$$

Then, from $2p = p_1 + p_2$, we are able to conclude that

$$2p - p^2 \leq p_1 + p_2 - p_1 p_2.$$

Using Lemma 2 ($\alpha_1 = \alpha_2 = 0.5$), it can be shown thhat

$$X_1 + X_2 <_{\mathrm{cx}} Y_1 + Y_2.$$

That is, the clrf of X_1 and X_2 is less than or equal to that of $Y_1 + Y_2$.

2) Let $F_n^{(p_1,\ldots,p_n)}(y)$ represent the clrf of n heterogeneous Bernoulli sources. The ith Bernoulli source X_i has an activity parameter p_i and a peak cell rate pcr. Let $F_n^{(p)}(y)$ represent the clrf of n homogeneous Bernoulli sources where each Bernoulli source has an activity parameter $p = (p_1 + \cdots + p_n)/n$ and a peak cell rate pcr. Suppose $F_n^{(p_1,\ldots,p_n)}(y) \leq F_n^{(p)}(y)$ holds for the case when $n = k$, i.e.

$$F_k^{(p_1,\ldots,p_n)}(y) \leq F_k^{(p)}(y), \quad \text{for any } y \in R$$

where $p = (p_1 + \cdots + p_k)/k$. Let us consider the case when $n = k + 1$. It can be shown that

$$\begin{aligned}
&F_{k+1}^{(p_1,\ldots,p_{k+1})}(y) \\
&\quad = F_k^{(p_1,\ldots,p_k)} * f_1^{(p_{k+1})}(y) \\
&\quad \leq F_k^{(\frac{p_1+\cdots+p_k}{k})} * f_1^{(p_{k+1})}(y) \\
&\quad = \left[F_{k-1}^{(\frac{p_1+\cdots+p_k}{k})} * f_1^{(\frac{p_1+\cdots+p_k}{k})}\right] * f_1^{(p_{k+1})}(y) \\
&\quad = \left[F_{k-1}^{(\frac{p_1+\cdots+p_k}{k})} * f_1^{(p_{k+1})}\right] * f_1^{(\frac{p_1+\cdots+p_k}{k})}(y) \\
&\quad \leq F_k^{(\frac{(k-1)\frac{p_1+\cdots+p_n}{k}+p_{k+1}}{k})} * f_1^{(\frac{p_1+\cdots+p_k}{k})}(y) \\
&\quad \cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

Define a sequence a_i so that the above procedure can be expressed as

$$\begin{aligned}
&F_{k+1}^{(p_1,\ldots,p_{k+1})}(y) \\
&\quad \leq F_k^{(a_1)} * f_1^{(a_0)}(y) \\
&\quad \cdots\cdots\cdots\cdots\cdots\cdots \\
&\quad \leq F_k^{(a_i)} * f_1^{(a_{i-1})}(y) \\
&\quad \cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

It can be shown that

$$\begin{aligned}
a_i &= \frac{(i-1)a_{i-1} + a_{i-2}}{i} \\
a_0 &= p_{k+1} \quad \text{and} \quad a_1 = \frac{p_1 + \cdots + p_k}{k}.
\end{aligned}$$

Solving for a_i, when $i \geq 2$, we obtain

$$a_i = \frac{p_1 + \cdots + p_k}{k} - \frac{1 - \left(-\frac{1}{k}\right)^{i-1}}{k+1} \times \left(\frac{p_1 + \cdots + p_k}{k} - p_{k+1}\right).$$

Thus, it can be shown that

$$\lim_{i\to\infty} a_i = \lim_{i\to\infty} a_{i-1} = \frac{p_1 + \cdots + p_{k+1}}{k+1}.$$

This means that, when the above process goes on and on, the activity parameters of the n homogeneous Bernoulli sources and the single Bernoulli source in the above equations will converge to $((p_1 + \cdots + p_{k+1})/(k+1))$. So it can be concluded that

$$\begin{aligned}
&F_{k+1}^{(p_1,\ldots,p_{k+1})}(y) \\
&\quad \leq F_k^{(\frac{p_1+\cdots+p_{k+1}}{k+1})} * f_1^{(\frac{p_1+\cdots+p_{k+1}}{k+1})}(y) \\
&\quad = F_{k+1}^{(\frac{p_1+\cdots+p_{k+1}}{k+1})}(y).
\end{aligned}$$

Therefore, from the supposition that $F_n^{(p)}(y) \geq F_n^{(p_1,\ldots,p_n)}(y)$ holds for the case when $n = k$, we derive the conclusion that the inequality holds for the case when $n = k + 1$.

Combining step 1 and step 2, it is concluded that

$$F_n^{(p)}(y) \geq F_n^{(p_1,\ldots,p_n)}(y), \quad \text{for any } y \in R$$

holds for any n. ■

APPENDIX B

In this appendix, we shall prove Theorem 3. First we introduce an important lemma which will be used in the proof of Theorem 3.

Lemma 3: Let X_1, X_2 be two independent ON–OFF sources with peak cell rates $\alpha_1 \times \mathrm{pcr}, (1-\alpha_1) \times \mathrm{pcr}$, and activity parameters p_1, p_2, respectively, where $0.5 \leq \alpha_1 < 1$. Let α_2 be any value satisfying $\alpha_1 \leq \alpha_2 < 1$. Then there exists two independent ON–OFF sources Y_1, Y_2 with peak cell rates $\alpha_2 \times \mathrm{pcr}, (1-\alpha_2) \times \mathrm{pcr}$, and $E(Y_1+Y_2) = E(X_1+X_2)$, such that $X_1 + X_2 <_{\mathrm{cv}} Y_1 + Y_2$.

See [30, Ch. 3] for a proof.

This lemma is used in proving Theorem 3 which is on heterogeneous ON–OFF sources. We shall now begin the proof of Theorem 3.

Proof: Theorem 3 is proved using the induction method.

1) Let us consider the case when $n = 1$, then $m = 1$. We must show that $X_1 <_{\mathrm{cx}} Y_1$, where X_1 is an ON–OFF source with peak cell rate pcr_1 and mean cell rate mcr_1, and Y_1 is an ON–OFF source with peak cell rate $\mathrm{pcr} \geq \mathrm{pcr}_1$ and mean cell rate mcr_1. This in fact is a direct application of Theorem 2.

2) Supposing inequality (8) holds for the case when $n = k$, we must show that inequality (8) also holds for the case when $n = k+1$.

First we point out that, since $\max\{\mathrm{pcr}_1, \ldots, \mathrm{pcr}_k\} \leq \max\{\mathrm{pcr}_1, \ldots, \mathrm{pcr}_{k+1}\}$, from our supposition that inequality (8) holds for $n = k$ and $\mathrm{pcr} \geq \max\{\mathrm{pcr}_1, \ldots, \mathrm{pcr}_k\}$, it naturally follows that it also holds for $n = k$ and $\mathrm{pcr} \geq \max\{\mathrm{pcr}_1, \ldots, \mathrm{pcr}_{k+1}\}$. Now let us consider the two ON–OFF source X_k and X_{k+1}, we shall consider the following two cases.

a) When $\mathrm{pcr}_k + \mathrm{pcr}_{k+1} \leq \mathrm{pcr}$, using Proposition 1, it can be shown that

$$X_k + X_{k+1} <_{\mathrm{cv}} Z$$

where Z is an ON–OFF source with peak cell rate $\mathrm{pcr}_Z = \mathrm{pcr}_k + \mathrm{pcr}_{k+1}$ and mean cell rate $\mathrm{mcr}_Z = \mathrm{mcr}_k + \mathrm{mcr}_{k+1}$. Then, applying our supposition for k ON–OFF sources $X_1, \ldots, X_{k-1}, Z$, it can be shown that inequality (8) holds for the $k+1$ ON–OFF sources $X_1, \ldots, X_{k+1}$, i.e.,

$$\begin{aligned} X_1 + \cdots + X_{k-1} + X_k + X_{k+1} &\\ <_{\mathrm{cv}} X_1 + \cdots + X_{k-1} + Z &\\ <_{\mathrm{cv}} Y_1 + \cdots + Y_{m_{k+1}} & \end{aligned}$$

where

$$m_{k+1} = \left\lceil \frac{\sum_{i=1}^{k-1} \mathrm{pcr}_i + \mathrm{pcr}_Z}{\mathrm{pcr}} \right\rceil = \left\lceil \frac{\sum_{i=1}^{k+1} \mathrm{pcr}_i}{\mathrm{pcr}} \right\rceil$$

and $Y_{i(i=1,\ldots,m_{k+1})}$ is an independent ON–OFF source with peak cell rate pcr and mean cell rate $\mathrm{mcr} = ((\sum_{i=1}^{k+1} \mathrm{mcr}_i)/(m_{k+1}))$.

b) When $\mathrm{pcr}_k + \mathrm{pcr}_{k+1} > \mathrm{pcr}$, using Lemma 3, we are able to find two independent ON–OFF source Z_1 and Z_2, Z_1 has a peak cell rate $\mathrm{pcr}_{Z_1} = \mathrm{pcr}_k + \mathrm{pcr}_{k+1} - \mathrm{pcr}$ and a mean cell rate mcr_{Z_1}, Z_2 has peak cell rate pcr and mean cell rate $\mathrm{mcr}_{Z_2}, \mathrm{mcr}_{Z_1} + \mathrm{mcr}_{Z_2} = \mathrm{mcr}_k + \mathrm{mcr}_{k+1}$, such that

$$X_k + X_{k+1} <_{\mathrm{cv}} Z_1 + Z_2.$$

Here $\mathrm{pcr} \geq \max\{\mathrm{pcr}_k, \mathrm{pcr}_{k+1}\}$ is equivalent to the condition in Lemma 3 that $\alpha_1 \leq \alpha_2$.

From $\mathrm{pcr} \geq \max\{\mathrm{pcr}_1, \ldots, \mathrm{pcr}_{k+1}\}$, it is guaranteed that $\mathrm{pcr} \geq \mathrm{pcr}_{Z_1}$. Then apply our supposition for the k independent ON–OFF sources $X_1, \ldots, X_{k-1}, Z_1$, we obtain

$$X_1 + \cdots + X_{k-1} + Z_1 <_{\mathrm{cv}} \lambda_1 + \cdots + \lambda_{m'_k}$$

where

$$\begin{aligned} m'_k &= \left\lceil \frac{\sum_{i=1}^{k-1} \mathrm{pcr}_i + \mathrm{pcr}_{Z_1}}{\mathrm{pcr}} \right\rceil \\ &= \left\lceil \frac{\sum_{i=1}^{k+1} \mathrm{pcr}_i - \mathrm{pcr}}{\mathrm{pcr}} \right\rceil \\ &= \left\lceil \frac{\sum_{i=1}^{k+1} \mathrm{pcr}_i}{\mathrm{pcr}} \right\rceil - 1 \end{aligned}$$

and $\lambda_{i(i=1,\ldots,m'_k)}$ is an independent ON–OFF source with peak cell rate pcr and mean cell rate

$$\mathrm{mcr}_\lambda = \frac{\sum_{i=1}^{k-1} \mathrm{pcr}_i + \mathrm{mcr}_{Z_1}}{m'_k}.$$

Here we note that $\lambda_{i(i=1,\ldots,m'_k)}$ and Z_2 have the same peak cell rate pcr, so, using Theorem 1, it can be shown that

$$\lambda_1 + \cdots + \lambda_{m'_k} + Z_2 <_{\mathrm{cv}} Y_1 + \cdots + Y_{m_{k+1}}$$

where

$$m_{k+1} = m'_k + 1 = \left\lceil \frac{\sum_{i=1}^{k+1} \mathrm{pcr}_i}{\mathrm{pcr}} \right\rceil$$

and $Y_{i(i=1,\ldots,m_{k+1})}$ is an independent ON–OFF source with peak cell rate pcr and mean cell rate

$$\begin{aligned} \mathrm{mcr} &= \frac{\sum_{i=1}^{k-1} \mathrm{pcr}_i + \mathrm{mcr}_{Z_1} + \mathrm{mcr}_{Z_2}}{m_{k+1}} \\ &= \frac{\sum_{i=1}^{k+1} \mathrm{pcr}_i}{m_{k+1}}. \end{aligned}$$

So, from the supposition that inequality (8) holds for $n = k$, we arrive at the conclusion that it should also holds for $n = k+1$.

Combining 1 and 2, we conclude that (8) holds for all n. ■

ACKNOWLEDGMENT

The authors wish to thank G. de Veciana at the University of Texas at Austin for his advice on using the stochastic ordering theory for our theoretical analysis and some other helpful comments. They are grateful to W. Feng of The Ohio State University for making his video traces available for simulation purpose, and to J. B. Micheel of Waikato Applied Network Dynamics group at Waikato University for making his traffic measurement data available for traffic analysis. They would also like to thank the anonymous reviewers for their helpful comments on the earlier draft of this paper.

REFERENCES

[1] J. Sugih, P. B. Danzig, S. J. Shenker, and Z. Lixia, "A measurement-based admission control algorithm for integrated service packet networks," *IEEE/ACM Trans. Networking*, vol. 5, pp. 56–69, Feb. 1997.

[2] D. Mitra, M. I. Reiman, and J. Wang, "Robust dynamic admission control for unified cell and call QoS in statistical multiplexers," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 692–707, 1998.

[3] C. Hao and L. Sanqi, "Fast algorithms for measurement-based traffic modeling," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 612–626, 1998.

[4] J. T. Lewis, R. Russell, F. Toomey, B. McGurk, S. Crosby, and I. Leslie, "Practical connection admission control for atm networks based on on-line measurements," *Comput. Commun.*, vol. 21, no. 17, pp. 1585–1596, 1998.

[5] K. Shiomoto and S.-I. Chaki, "Adaptive connection admission control using real-time traffic measurements in ATM networks," *IEICE Trans. Commun.*, vol. E78-B, no. 4, pp. 458–464, 1995.

[6] R. Bolla, F. Davoli, and M. Marchese, "Bandwidth allocation and admission control in atm networks with service separation," *IEEE Commun. Mag.*, vol. 35, pp. 130–137, May 1997.

[7] R. J. Gibbens, F. P. kelly, and P. B. Key, "A decision-theoretic approach to call admission control in atm networks," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1101–1114, 1995.

[8] S. Lee and J. Song, "A measurement-based admission control algorithm using variable-sized window in atm networks," *Computer Commun.*, vol. 21, no. 2, pp. 171–178, 1998.

[9] M. Fontaine and D. G. Smith, "Bandwidth allocation and connection admission control in atm networks," *Electron. Commun. Eng. J.*, vol. 8, no. 4, pp. 156–164, 1996.

[10] A. E. Eckberg, "B-ISDN/ATM traffic and congestion control," *IEEE Network Mag.*, vol. 6, pp. 28–37, 1992.

[11] C. Rasmussen, J. H. Sorensen, K. S. Kvols, and S. B. Jacobsen, "Source-independent call acceptance procedures in atm networks," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 351–358, 1991.

[12] T.-H. Lee, K.-C. Lai, and S.-T. Duann, "Design of a real-time call admission controller for atm networks," *IEEE/ACM Trans. Networking*, vol. 4, no. 5, pp. 758–765, 1996.

[13] M. Reisslein, K. W. Ross, and S. Rajagopal, "Guaranteeing statistical QoS to regulated traffic: The single node case," in *IEEE Infocom'1999*, Apr. 1999.

[14] J. Y. Hui, "Resource allocation for broadband networks," *IEEE J. Select. Areas Commun.*, vol. 6, pp. 1598–1608, Sept. 1988.

[15] R. Szekli, *Stochastic Ordering and Dependence in Applied Probability*. New York: Springer-Verlag, 1995.

[16] S. M. Ross, *Stochastic Processes*. New York: Wiley, 1983.

[17] Z.-L. Zhang, J. Jurose, J. Salehi, and D. Towsley. (1996) Smoothing, Statistical Multiplexing and Call Admission Control for Stored Video. Dept. of Computer Science, University of Massachusetts, Amherst, MA. [Online]. Available: gaia.cs.umass.edu in pub/Zhan96:Smoothing.ps.gz.

[18] B. Maglaris, D. Anastassiou, P. Sen, G. Kaelsson, and J. D. Robbins, "Performance models of statistical multiplexing in packet video communications," *IEEE Trans. Commun.*, vol. 36, pp. 834–844, July 1988.

[19] T. Murase, H. Suzuki, S. Sato, and T. Takeuchi, "A call admission control scheme for atm networks using a simple quality estimate," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 1461–1470, Sept. 1991.

[20] N. G. Bean, "Statistical Multiplexing in Broadband Communication," Ph.D. thesis, Univ. Cambridge, Cambridge, 1993.

[21] R. Griffiths and P. Key, "Adaptive call admission control in ATM networks," in *Proc. 14th Int. Teletraffic Cong.-ITC 14*, June 1994.

[22] M. Grossglauser and D. N. C. Tse, "A framework for robust measurement-based admission control," *IEEE/ACM Trans. Networking*, vol. 7, pp. 293–309, June 1999.

[23] K. Shiomoto, S. Chaki, and N. Yamanaka, "A simple bandwidth management strategy based on measurements of instantaneous virtual pacth utilization in atm networks," *IEEE/ACM Trans. Networking*, vol. 6, pp. 625–634, Oct. 1998.

[24] Z. Dziong, M. Juda, and L. G. Mason, "A framework for bandwidth management in ATM networks—Aggregate equivalent bandwidth estimation approach," *IEEE/ACM TRans. Networking*, vol. 5, pp. 134–147, Feb. 1997.

[25] M. Grossglauser and D. N. C. Tse, "A time-scale decomposition approach to measurement-based admission control," *Proc. IEEE INFOCOM*, 1999.

[26] S. M. Srinidhi, W. H. Thesling, and V. K. Konangi, "An adaptive scheme for admission control in atm networks," *Computer Networks and ISDN Systems*, vol. 29, pp. 569–582, 1997.

[27] H. Yamada and S. Sumita, "A traffic measurement method and its application for cell loss probability estimation in atm networks," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 315–324, June 1991.

[28] M. E. Crovella and A. Bestavros, "Self-similarity in world wide web traffic: Evidence and possible causes," *IEEE/ACM Trans. Networking*, vol. 5, pp. 835–846, Dec. 1997.

[29] W. E. Leland, M. S. Taqqu, W. R. Willinger, and D. V. Wilson, "On the self-similar nature of ethernet traffic (extended version)," *IEEE/ACM Trans. Networking*, vol. 2, pp. 1–15, Feb. 1994.

[30] G. Mao, "Statistical Multiplexing and Connection Admission Control in ATM Networks," Ph.D. Thesis, School of Engineering and Mathematics, Edith Cowan University, Apr. 2001.

[31] J. W. Roberts, "Variable-bit-rate traffic control in B-ISDN," *IEEE Commun. Mag.*, vol. 29, pp. 50–56, Sept. 1991.

[32] R. J. Gibbens and P. J. Hunt, "Effective bandwidths for multi-type UAS channel," *Queueing Syst.*, vol. 9, no. 1-2, pp. 17–28, 1991.

[33] R. Guerin, H. Ahmadi, and M. Naghshineh, "Equivalent capacity and its application to bandwidth application in high-speed networks," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 968–981, 1991.

[34] Video-on-Demand Services: Efficient Transportation and Decompression of Variable Bit Rate Video, W.-C. Feng. (1996, Apr.). [Online]. Available: http://www.cis.ohio-state.edu/ wuchi/.

[35] M. Grossglauser and J.-C. Bolot, "On the relevance of long-range dependence in network traffic," *IEEE/ACM Trans. Networking*, vol. 7, pp. 629–640, Oct. 1999.

[36] B. K. Ryu and A. Elwalid, "The importance of long-range dependence of VBR video traffic in ATM traffic engineering: Myths and realities," *Comput. Commun. Rev.*, vol. 24, no. 4, Oct. 1996.

[37] Y. Kim and S.-Q. Li, "Timescale of interest in traffic measurement for link bandwidth allocation design," *Proc. IEEE INFOCOM'96*, Apr. 1996.

[38] D. P. Heyman and T. V. Lakshman, "What are the implications of long-range dependence for VBR-video traffic engineering?," *IEEE/ACM Trans. Networking*, vol. 4, pp. 301–317, June 1996.

Guoqiang Mao (S'98) was born in Yichang, China, in April 1974. He received the B.S. degree in electrical engineering from Hubei Institute of Technology, China, in 1995, the M.S. degree in instrumentation engineering from Southeast University, China, in 1998, and the Ph.D. degree in telecommunication networks from Edith Cowan University, Joondalup, Australia, in 2001.

He is now a Senior Engineer at Intelligent Pixel Incorporation, Australia. His research interests include telecommunication system modeling, traffic analysis and modeling, connection admission control and optical communications.

Daryoush Habibi (M'95) received the B.S. degree in electrical engineering and the Ph.D. degree in telecommunications from the University of Tasmania, Australia, in 1989 and 1993, respectively.

During his professional career, he has worked with a number of industries and academic institutions, including Telstra Research Laboratories, Flinders University, Edith Cowan University, and Intelligent Pixels Incorporation, where he is currently the Engineering Vice-President. His interests include system modeling and performance evaluation for communication systems and networks. Another prime area of his interest is system development for optical switching and networking.