# Characterization of Delay-Sensitive Traffic

Tsern-Huei Lee, *Member, IEEE,* and Kuen-Chu Lai

***Abstract*—Resource allocation is necessary for a network which guarantees quality of service (QoS). In this paper we first present a definition for a traffic stream to be burstier than another traffic stream. The definition is based on the loss probability of a bufferless multiplexer and thus is appropriate for delay-sensitive traffic which cannot tolerate queueing delay caused by buffering. An optimum quantization algorithm is then derived for source characterization. The optimally quantized version achieves minimum loss rate for all possible allocated bandwidths under the condition that it is burstier than the real traffic. The quantized source is called a pseudosource and can be used by the network for resource allocation. Some numerical examples are studied. Results show that, for a bufferless multiplexer, the allocated bandwidth based on optimally quantized pseudosources is only slightly greater than the minimum bandwidth required to meet the requested QoS.**

***Index Terms*— Bufferless multiplexer, burstiness, delay-sensitive traffic, optimum quantization, variability ordering.**

## I. INTRODUCTION

RESOURCE allocation is necessary for a network [such as the broad-band Integrated Services Digital Network (ISDN)] which guarantees quality of service (QoS). Of course, the network would like to avoid overallocating resources to increase system utilization. To achieve this, traffic sources must be characterized as accurately as possible.

Burstiness is considered an important aspect of traffic characteristics. One possible definition of *"traffic stream $X(t)$ is burstier than traffic stream $Y(t)$"* is that the steady-state queue length of a stable deterministic single-server queue with arrival process $X(t)$ is stochastically larger than the steady-state queue length of the same queue but with arrival process $Y(t)$ [2]. In other words, $X(t)$ is burstier than $Y(t)$ if $F_X(x) \leq F_Y(x)$ for $x \geq 0$, where $F_X(x)$ and $F_Y(x)$ are the queue-length distributions of the single-server queue when arrival process is $X(t)$ or $Y(t)$, respectively. More definitions of burstiness can be found in [3].

In this paper we provide an alternative definition, based on the loss probability of a bufferless multiplexer. The definition is aimed at characterizing delay-sensitive traffic streams which do not allow delay caused by buffering. This definition is utilized to characterize (or quantize) traffic streams for networks which provide only finitely many bit rates for users to describe their traffic. To guarantee QoS, the quantized version has to be burstier than the real traffic. We present an optimum quantization algorithm which yields a quantized version that is burstier than the real source and achieves minimum loss rate for all possible allocated bandwidths. The quantized version, called a pseudosource, is used to represent the real source and is considered by the network for bandwidth allocation.

Manuscript received April 24, 1997; revised January 28, 1998 and March 6, 1998; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor S. Low. This work was supported in part by the National Science Council, Taiwan, R.O.C., under Contract NSC-86-2213-E009-069.

T.-H. Lee is with the Department of Communication Engineering and the Center for Telecommunications Research, National Chiao Tung University, Hsinchu 30050, Taiwan, R.O.C. (e-mail: thlee@atm.cm.nctu.eud.tw).

K.-C. Lai is with the Voice of Han Broadcasting Network, Taipei, Taiwan, R.O.C. (e-mail: kclainet @ms12.hinet.net).

Publisher Item Identifier S 1063-6692(98)05134-6.

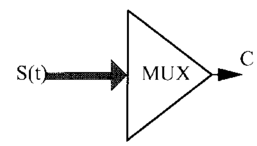


Fig. 1. A bufferless multiplexer with link capacity $C$.

In [6] the throughput loss due to bandwidth quantization was evaluated assuming that peak rate is the only parameter for traffic description. The results are applicable to multirate circuit switching networks. For packet switching networks, system utilization can be significantly reduced without taking into account statistical multiplexing. The main focus of [6] is to choose the quantization rates to minimize throughput loss. Here, we assume that the rates provided by a network are fixed (which is often the case) and derive an optimum source traffic quantization.

The rest of this paper is organized as follows. In Section II we describe the investigated multiplexer system. In Section III we give a definition of a traffic stream to be burstier than another traffic stream and prove some properties. In Section IV an optimum quantization algorithm is presented. Numerical examples are studied in Section V. Results show that, for a bufferless multiplexer, the bandwidth allocated based on pseudosources is only slightly greater than the minimum bandwidth required to guarantee the requested QoS. A conclusion is finally drawn in Section VI.

## II. THE INVESTIGATED SYSTEM

Fig. 1 shows a multiplexer with link capacity $C$. The traffic which arrives at the multiplexer is denoted by $S(t)$. Notice that $S(t)$ represents the aggregate traffic generated by existing sources. In this paper all traffic sources are assumed to be nonnegative, bounded, stationary, and ergodic. Moreover, we assume that all sources generate delay-sensitive traffic independently and that the buffer size in the multiplexer is small so that its effect can be neglected.

Let $S$ denote a random variable whose distribution is identical to that of $S(t)$ for all $t$. For the bufferless multiplexer, the loss rate $\mathrm{Loss}(T)$ in $[0, T]$ is given by

$$\mathrm{Loss}(T) = \frac{1}{T}\int_0^T (S(t) - C)^+ \, dt$$

where $(x)^+ = \max\{0, x\}$. Since all sources are assumed to be ergodic, the average loss rate, denoted by $L_S(C)$, can be computed as follows:

$$L_S(C) = \lim_{T\to\infty} \text{Loss}(T) = E(S - C)^+.$$

Furthermore, the loss probability is given by $(E(S - C)^+)/(E(S))$.

## III. Burstiness of Traffic Sources

Let $X(t)$ and $Y(t)$ be two traffic streams that are independent of $S(t)$. With the assumptions of stationarity and ergodicity, to study the loss probability of the bufferless multiplexer, it suffices to use two nonnegative bounded random variables $X$ and $Y$ to represent $X(t)$ and $Y(t)$, respectively.

We define "$X$ is burstier than $Y$" as follows.

*Definition 1:* We say traffic stream $X$ is burstier than traffic stream $Y$, denoted by $X \geq_B Y$, if and only if (iff) adding $X$ into the bufferless multiplexer with any existing traffic $S$ (which is independent of $X$ and $Y$) results in a greater loss probability than adding $Y$.

Mathematically, $X \geq_B Y$ iff

$$\frac{E(S+X-C)^+}{E(S+X)} \geq \frac{E(S+Y-C)^+}{E(S+Y)} \tag{1}$$

for any nonnegative bounded random variable $S$ which is independent of $X$ and $Y$.

We assume for the rest of this paper that all given random variables are independent. We now state and prove some properties regarding loss rate and the above definition.

*Property 1:* If $X_i \geq_B Y_i$, $i = 1, 2, \cdots, N$, then $\sum_{i=1}^N X_i \geq_B \sum_{i=1}^N Y_i$.

*Proof:* According to Definition 1, we get $X + W \geq_B Y + W$ for any traffic $W$ if $X \geq_B Y$. Therefore, we have

$$\begin{aligned} X_1 + \sum_{i=2}^N X_i &\geq_B Y_1 + \sum_{i=2}^N X_i \\ &\geq_B Y_1 + Y_2 + \sum_{i=3}^N X_i \geq_B \cdots \geq_B \sum_{i=1}^N Y_i. \end{aligned}$$

This completes the proof of Property 1. □

*Property 2:* If $X \geq_B Y$, then $E(X) \geq E(Y)$ and $L_X(c) \geq L_Y(c)$ for all $c$, $0 \leq c \leq C$.

*Proof:* Since $X \geq_B Y$, letting $S = C - c$, one gets $((E(X-c)^+)/(E(X)+C-c)) \geq ((E(Y-c)^+)/(E(Y)+C-c))$ for all $c$, $0 \leq c \leq C$. Similarly, letting $S = C$, one has $((E(X))/(E(X+C))) \geq ((E(Y))/(E(Y+C)))$, which implies $E(X) \geq E(Y)$. Combining the above results, we get $E(X-c)^+ \geq E(Y-c)^+$ for all $c$, $0 \leq c \leq C$. This completes the proof of Property 2. □

For convenience, we shall use $X \geq_R Y$ to denote $L_X(c) \geq L_Y(c)$ for all $c$, $0 \leq c \leq C$. Notice that, for $X \leq C$ and $Y \leq C$, the relation $\geq_R$ is identical to the variabilty (or convex) ordering, and $X \geq_R Y$ means that $X$ is more variable than $Y$ [7].

*Property 3:* If $E(X) = E(Y)$ and $X \geq_R Y$, then $X \geq_B Y$.

*Proof:* This property is proven if one can show that $X \geq_R Y$ implies $E[S+X-C]^+ \geq E[S+Y-C]^+$ for all traffic $S$. Since $X \geq_R Y$ implies $E(X-c)^+ \geq E(Y-c)^+$ for all $c$, $c \leq C$, we have

$$\begin{aligned} E(S+X-C)^+ &= E\{E[(S+X-C)^+ \mid S]\} \\ &\geq E\{E[(S+Y-C)^+ \mid S]\} \\ &= E(S+Y-C)^+. \end{aligned}$$

Therefore, $E(X) = E(Y)$ and $X \geq_R Y$ implies $X \geq_B Y$. □

*Property 4:* There does not exist a total ordering definition of burstiness such that $X$ *is burstier than* $Y$ implies $X \geq_B Y$.

*Proof:* Since $X \geq_B Y$ implies $X \geq_R Y$, this property is proven if one can show $X \geq_R Y$ is not a total ordering. This can be done with a counterexample. Let $X$ and $Y$ be two-state sources with probability mass functions

$$P_X(i) = \begin{cases} 0.5, & \text{if } i = 0 \\ 0.5, & \text{if } i = \frac{3}{4}C \end{cases}$$

$$P_Y(i) = \begin{cases} \frac{4}{5}, & \text{if } i = \frac{1}{4}C \\ \frac{1}{5}, & \text{if } i = \frac{5}{4}C \end{cases}$$

respectively. As a result, we have $E(X - \frac{1}{4}C)^+ = \frac{1}{4}C > E(Y-C)^+ = \frac{1}{5}C$; however, $E(X-C)^+ = 0 < E(Y - C)^+ = \frac{1}{20}C$. In other words, neither $X \geq_R Y$ nor $Y \geq_R X$ holds. This completes the proof of Property 4. □

Property 4 says that any definition of burstiness in terms of a real number, say, the ratio of peak rate to average rate [3], may not be appropriate.

*Property 5:* Suppose $X \geq_B Y$, $X \leq C$, and $Y \leq C$. If there exists a $C_0$, $0 \leq C_0 < C$, such that $L_X(C_0) > L_Y(C_0)$, then $E(f(X)) > E(f(Y))$ for any increasing strictly convex function $f$ on $[0, \infty)$ for which $f''(x)$ exists and is continuous.

*Proof:* Let $F(x)$ denote the distribution function of $X$ and $\bar{F}(x) = 1 - F(x)$. It can be shown (see, for example, [7, proof of Proposition 8.5.1]) that

$$\begin{aligned} E[f(X)] &= \int_0^\infty f''(c) L_X(c)\, dc + f(0) + f'(0) L_X(0) \\ &= \int_0^C f''(c) L_X(c)\, dc + f(0) + f'(0) L_X(0). \end{aligned}$$

Similar identity can be obtained for random variable $Y$. Therefore, we have $E[f(X)] - E[f(Y)] = f'(0)[L_X(0) - L_Y(0)] + \int_0^C f''(c)[L_X(c) - L_Y(c)]\, dc$. The fact that $f(x)$ is an increasing strictly convex function implies $f'(x) > 0$ and $f''(x) > 0$ for all $x$. Besides, $X \geq_B Y$ implies $X \geq_R Y$. As a result, we get

$$E[f(X)] - E[f(Y)] \geq \int_0^C f''(c)[L_X(c) - L_Y(c)]\, dc.$$

Since $L_X(C_0) > L_Y(C_0)$ for some $C_0$, $0 \leq C_0 < C$, we conclude that $E[f(X)] - E[f(Y)] > 0$ because both $L_X(c)$ and $L_Y(c)$ are continuous. This completes the proof of Property 5. □

Property 5 implies the commonly used moment-matching technique [5] in approximating a traffic source may not be appropriate for delay-sensitive traffic. The reason is that $f(x) = x^n$, $n \geq 2$ is a twice-differentiable increasing strictly convex function for $x > 0$ and $f(0) = 0$. Since the average rate is not changed for the moment-matching technique, the loss probability of the approximating source is not guaranteed to be an upper bound of that of the approximated source. For example, suppose we want to approximate a four-state source $Y$ with an on–off source $X$. Let $\mathbf{q} = [q_1 \;\; q_2 \;\; q_3 \;\; q_4]$ and $\mathbf{m} = [m_1 \;\; m_2 \;\; m_3 \;\; m_4]$ denote, respectively, the stationary probability vector and the bit-rate vector of $Y$. Also, let $\mathbf{p} = [p_1 \;\; p_2]$ and $\mathbf{M} = [0 \;\; M]$ be those of $X$. Assume $\mathbf{q} = [0.4 \;\; 0.3 \;\; 0.2 \;\; 0.1]$ and $\mathbf{m} = [1 \;\; 2 \;\; 3 \;\; 4]$ Mb/s. The conditions $E(X^n) = E(Y^n)$, $1 \leq n \leq 2$ result in $\mathbf{p} = [0.2 \;\; 0.8]$ and $\mathbf{M} = [0 \;\; 2.5]$ Mb/s. If $c = 2.2$ Mb/s $< C$, then we have $L_X(2.2) = 0.24$ Mb/s $< L_Y(2.2) = 0.34$ Mb/s.

## IV. Quantization of Traffic Sources

A network in the real world is likely to provide only finitely many bit rates for users to describe their traffic characteristics. In this case, one has to quantize traffic sources. In this section we present an optimal quantization algorithm. The quantization is optimum in the sense of achieving minimum loss rate for all possible allocated bandwidths, under the condition that the quantized version is burstier than the real source.

Assume there are $N + 1$ quantized bit rates denoted by $M_0, M_1, \cdots$, and $M_N$ such that $M_i < M_j$ if $i < j$ and $M_0 = 0$. The quantized source has $N+1$ states. A quantization is an assignment of the stationary probability vector $\mathbf{p} = [p_0 \;\; p_1 \;\; p_2 \;\; \cdots \;\; p_N]$ to the quantized source. Therefore, for convenience, we call a probability vector $\mathbf{p}$ a quantization. Of course, some entries of $\mathbf{p}$ can be zero. It should be noted that quantization does not alter characteristics of the real traffic. The quantized source is a pseudosource considered by the network only for resource allocation.

Let $Y$ denote the traffic generated by the source. We assume that $Y$ is nonnegative and $\mathrm{Prob}(Y > M_N) = 0$. Also, let $\{\mathbf{p}, \mathbf{M}, Z(\mathbf{p}), \mathbf{N+1}\}$ represent a quantized pseudosource, where $\mathbf{p}$ denotes the quantization and $Z(\mathbf{p})$ denotes the traffic generated by the pseudosource. Since the pseudosource is used for resource allocation, it is required to be burstier than the real source. Therefore, we define a quantization $\mathbf{p}$ to be legal iff it satisfies

$$Z(\mathbf{p}) \geq_B Y$$

i.e., iff $Z(\mathbf{p})$ is burstier than $Y$. An optimum quantization could be defined as a legal quantization $\mathbf{p}^*$ which satisfies $Z(\mathbf{p}) \geq_B Z(\mathbf{p}^*) \geq_B Y$ for any legal quantization $\mathbf{p}$. However, because of the presence of a nonnegative random variable $S$ in (1), it is rather difficult to determine whether or not an optimum quantization exists under this sense and, if it does exist, how to determine it. Therefore, we modify the criterion and say that a legal quantization is optimum if it achieves $\min_{\mathbf{p}\in\Omega} E(Z(\mathbf{p}) - c)^+$ for all $c$, where $\Omega$ denotes the set of legal quantizations. The definition is given for all $c$ because the allocated bandwidth may vary in a real network. For example, the bandwidth allocated to a virtual path in an asynchronous transfer mode (ATM) network is likely to be time varying, depending on its carried traffic. It turns out that the optimum quantization $\mathbf{p}^*$ satisfies $E(Z(\mathbf{p}^*)) = E(Y)$. Therefore, according to Property 3, the optimally quantized version is burstizer than the original source (i.e., $Z(\mathbf{p}^*) \geq_B Y$). Besides, if $\hat{\mathbf{p}}$ exists, then $\hat{\mathbf{p}} = \mathbf{p}^*$ (see Theorem 2).

In the following we determine the optimum quantization of $Y$. Since $E(Z(\mathbf{p}) - c)^+ = 0$ for $c \geq M_N$ for any quantization $\mathbf{p}$, to determine the optimum quantization, it suffices to consider the region $c \in (-\infty, M_N]$.

*Lemma 1:* Let $y(c) = E[Y - c]^+$, then $y(c)$ is a decreasing convex function on $(-\infty, M_N]$. Moreover, $y(c)$ is linear on $(-\infty, 0]$.

*Proof:* The proof of Lemma 1 is simple and thus omitted. □

The following theorem shows the existence of a unique optimum quantization and how to determine it.

*Theorem 1:* Let $\mathbf{p}^* = [p_0^* \;\; p_1^* \;\; p_2^* \;\; \cdots \;\; p_N^*]$ be determined by

$$\sum_{k=0}^{N} p_k^*(M_k - M_i)^+ = E[Y - M_i]^+, \quad 0 \leq i \leq N-1$$

and $\sum_{i=0}^{N} p_i^* = 1$. Then $\mathbf{p}^*$ is the optimum quantization and $E[Z(\mathbf{p}^*)] = E[Y]$.

*Proof:* We shall first prove $\mathbf{p}^*$ is a quantization (i.e., a probability vector) by showing $p_i^* \geq 0$ for all $i$, $0 \leq i \leq N$. Let $y(c) = E[Y - c]^+$. It is clear that $f_{\mathbf{p}^*}(c) = \sum_{k=0}^{N} p_k^*(M_k - c)^+$ is a piecewise linear function and is linear on $(-\infty, 0]$ and $[M_i, M_{i+1}]$ for all $i$, $0 \leq i \leq N-1$. Moreover, the equations used to determine $\mathbf{p}^*$ imply $f_{\mathbf{p}^*}(M_i) = y(M_i)$ for all $i$, $0 \leq i \leq N$. In other words, $f_{\mathbf{p}^*}(c)$ passes through the point $(M_i, y(M_i))$, $0 \leq i \leq N$. For $c \in (-\infty, 0]$, we have

$$\begin{aligned} f_{\mathbf{p}^*}(c) &= \sum_{k=0}^{N} p_k^* M_k - c\sum_{k=0}^{N} p_k^* \\ &= f_{\mathbf{p}^*}(M_0) - c = y(M_0) - c = y(c) \end{aligned} \tag{2}$$

which implies that the modulus of the slope of $f_{\mathbf{p}^*}(c)$ is $\sum_{i=0}^{N} p_k = 1$. Consider the interval $c \in [M_i, M_{i+1}]$ for some $i \geq 0$. In this region we have

$$f_{\mathbf{p}^*}(c) = \sum_{k=i+1}^{N} p_k^*(M_k - c) = -c\sum_{k=i+1}^{N} p_k^* + \sum_{k=i+1}^{N} p_k^* M_k. \tag{3}$$

On the other hand, $f_{\mathbf{p}^*}(c)$ for $c \in [M_i, M_{i+1}]$ is linear and passes through the two points $(M_i, y(M_i))$ and $(M_{i+1}, y(M_{i+1}))$. Thus, it can be expressed as

$$f_{\mathbf{p}^*}(c) = (M_{i+1} - c)\frac{y(M_i) - y(M_{i+1})}{M_{i+1} - M_i} + y(M_{i+1}). \tag{4}$$

Comparing (3) and (4), we get

$$\sum_{k=i+1}^{N} p_k^* = \frac{y(M_i) - y(M_{i+1})}{M_{i+1} - M_i} \geq 0$$

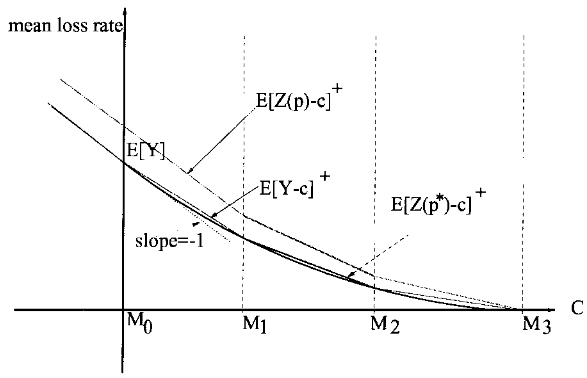


Fig. 2. Mean loss rate versus link capacity.

because, according to Lemma 1, $y(c)$ is a decreasing function. The special case when $i = N-1$ gives $p_N^* \geq 0$. Similarly, the modulus of the slope of the line segment of $f_{\mathbf{p}^*}(c)$ for $c \in [M_{i+1}, M_{i+2}]$ is $\sum_{k=i+2}^{N} p_k^*$. Since $y(c)$ is convex, it holds that $1 = \sum_{k=0}^{N} p_k^* \geq \sum_{k=i+1}^{N} p_k^* \geq \sum_{k=i+2}^{N} p_k^*$ for all $i$, $0 \leq i \leq N-2$. As a result, we get $p_i^* \geq 0$ for all $i$, $0 \leq i \leq N$. Therefore, the $\mathbf{p}^*$ determined by

$$\sum_{k=0}^{N} p_k^*(M_k - M_i)^+ = E[Y - M_i]^+, \qquad \text{for } 0 \leq i \leq N-1$$

$$\sum_{i=0}^{N} p_i^* = 1$$

is indeed a quantization. By substitution, we have $E[Z(\mathbf{p}^*)] = f_{\mathbf{p}^*}(0) = f_{\mathbf{p}^*}(M_0) = y(M_0) = E[Y]$. Since the line segment of $f_{\mathbf{p}^*}(c)$ coincides with $y(c)$ [see (2)] for $c \in (-\infty, 0]$ and lies above $y(c)$ for $c \in [M_i, M_{i+1}]$, $0 \leq i \leq N-1$, we conclude that $\mathbf{p}^*$ is a legal quantization.

We now prove that $\mathbf{p}^*$ is the unique optimum quantization. For any legal quantization $\mathbf{p}$, we must have $f_{\mathbf{p}}(M_i) \geq y(M_i)$ and $f_{\mathbf{p}}(M_{i+1}) \geq y(M_{i+1})$. Consequently, for $c \in [M_i, M_{i+1}]$, $0 \leq i \leq N-1$, the line segment of $f_{\mathbf{p}^*}(c)$ is below the line segment of $f_{\mathbf{p}}(c)$ if $f_{\mathbf{p}}(M_i) > y(M_i)$ or $f_{\mathbf{p}}(M_{i+1}) > y(M_{i+1})$, or both. Therefore, $\mathbf{p}^*$ is the unique optimum quantization. This completes the proof of Theorem 1. □

Theorem 1 gives the optimum quantization of any bounded traffic source to an $(N+1)$-state source. Notice that the mean of the real source is equal to the mean of the optimally quantized version. Fig. 2 shows an example of $f_{\mathbf{p}^*}(c)$ for $N = 3$.

*Theorem 2:* If there exists a legal quantization $\hat{\mathbf{p}}$ which satisfies $Z(\mathbf{p}) \geq_B Z(\hat{\mathbf{p}})$ for any legal quantization $\mathbf{p}$, then $\hat{\mathbf{p}} = \mathbf{p}^*$.

*Proof:* Assume that such a $\hat{\mathbf{p}}$ exists. According to Theorem 1, $\mathbf{p}^*$ is the unique legal quantization which satisfies $Z(\mathbf{p}) \geq_R Z(\mathbf{p}^*)$ for any legal quantization $\mathbf{p}$. Since $\hat{\mathbf{p}}$ is a legal quantization, we have $Z(\hat{\mathbf{p}}) \geq_R Z(\mathbf{p}^*)$. On the other hand, $Z(\mathbf{p}^*) \geq_B Z(\hat{\mathbf{p}})$ implies $Z(\mathbf{p}^*) \geq_R Z(\hat{\mathbf{p}})$, according to Property 2. Combining the above results, we get $E[Z(\hat{\mathbf{p}}) - c]^+ = E[Z(\mathbf{p}^*) - c]^+$ for all $c$, $0 \leq c \leq C$. Consequently, $Z(\mathbf{p}) \geq_R Z(\hat{\mathbf{p}})$ for any legal quantization $\mathbf{p}$. Since $\mathbf{p}^*$ is unique, we conclude that $\hat{\mathbf{p}} = \mathbf{p}^*$. This completes the proof of Theorem 2. □

In the following theorem we consider optimum quantization of $J$-state sources.

*Theorem 3:* Consider a $J$-state source described by $\{\mathbf{q}, \mathbf{m}, \mathbf{Y}, \mathbf{J}\}$. Let $\underline{m_i}$ be the largest quantization level that is smaller than or equal to $m_i$ and let $\bar{m}_i$ be the smallest quantization level that is larger than or equal to $m_i$. Let $Q = \{\underline{m}_0, \bar{m}_0, \underline{m}_1, \bar{m}_1, \cdots, \underline{m}_{J-1}, \bar{m}_{J-1}\}$. The optimum quantization $\mathbf{p}^*$ of the $J$-state source is given by

$$p_n^* = \frac{1}{M_n - M_{n-1}} \left[ \sum_{k=0}^{J-1} q_k (m_k - M_{n-1})^+ - \sum_{k=n+1}^{N} p_k^* (M_k - M_{n-1}) \right], \quad \text{if } n \in \{j > 0 : M_j \in Q\}$$

$$p_n^* = 0, \qquad \text{if } n \in \{j > 0 : M_j \notin Q\}$$

$$p_0^* = 1 - \sum_{n=1}^{N} p_n^*.$$

*Proof:* The first equation (i.e., $p_n^*$ for $n \in \{j > 0 : M_j \in Q\}$) is a direct result of Theorem 1. For example, the expression for $p_n^*$ can be derived from $E[Z(\mathbf{p}^*) - M_{n-1}]^+ = E[Y - M_{n-1}]^+$. Therefore, we need only prove the second equation, i.e., $p_n^* = 0$ if $n \in \{j > 0 : M_j \notin Q\}$. Assuming the set $\{j > 0 : M_j \notin Q\}$ is not empty, we have $\bar{m}_{j-1} \leq M_{n-1} < M_n < M_{n+1} \leq \underline{m}_j$ for some $j$. From Theorem 1 (with $i = n-1$, $n$, and $n+1$), we get

$$\sum_{k=j}^{J-1} q_k (m_k - M_{n-1}) = p_n^* (M_n - M_{n-1}) + \sum_{k=n+1}^{N} p_k^* (M_k - M_{n-1}) \tag{5}$$

$$\sum_{k=j}^{J-1} q_k (m_k - M_n) = \sum_{k=n+1}^{N} p_k^* (M_k - M_n) \tag{6}$$

$$\sum_{k=j}^{J-1} q_k (m_k - M_{n+1}) = \sum_{k=n+1}^{N} p_k^* (M_k - M_{n+1}). \tag{7}$$

Combining (5) and (6), one has $\sum_{k=j}^{J-1} q_k (M_n - M_{n-1}) = p_n^* (M_n - M_{n-1}) + \sum_{k=n+1}^{N} p_k^* (M_n - M_{n-1})$, which implies

$$p_n^* - \left( \sum_{k=j}^{J-1} q_k - \sum_{k=n+1}^{N} p_k^* \right) = 0. \tag{8}$$

Similarly, from (6) and (7), one has

$$\sum_{k=j}^{J-1} q_k = \sum_{k=n+1}^{N} p_k^*. \tag{9}$$

Finally, comparing (8) and (9), we have $p_n^* = 0$. This completes the proof of Theorem 3. □

A consequence of Theorem 3 is that the optimally quantized version of a $J$-state source can have as many as $2J+1$ states. Since the processing complexity in resource allocation may depend on (or even be proportional to) the number of states of the quantized version, one might wish to reduce it with only a little sacrifice in utilization. The following theorem states a suboptimum quantization algorithm that results in a pseudosource which has at most $J+1$ states. The suboptimum quantization algorithm is optimum if the quantization levels are restricted to $M_0$, $\bar{m}_0$, $\bar{m}_1, \cdots$, and $\bar{m}_{J-1}$.

*Theorem 4:* For a $J$-state source described by $\{\mathbf{q}, \mathbf{m}, \mathbf{Y}, \mathbf{J}\}$, if the allowed quantization levels are restricted to $M_0$, $\bar{m}_0$, $\bar{m}_1, \cdots$, and $\bar{m}_{J-1}$, then the optimum quantization $\mathbf{p}^*$ is determined by

$$\sum_{k=0}^{J-1} p_{k+1}^*[\bar{m}_k - \bar{m}_i]^+ = \sum_{k=0}^{J-1} q_k[m_k - \bar{m}_i]^+, \quad \text{for } 0 \le i \le J-2$$

$$\sum_{k=0}^{J-1} p_{k+1}^*[\bar{m}_k - M_0]^+ = \sum_{k=0}^{J-1} q_k[m_k - M_0]^+$$

$$p_0^* = 1 - \sum_{i=1}^{J-1} p_i^*.$$

If $\bar{m}_i \neq \bar{m}_j$ for $i \neq j$, then the result is

$$p_J^* = \frac{m_{J-1} - \bar{m}_{J-2}}{\bar{m}_{J-1} - \bar{m}_{J-2}} q_{J-1}$$

$$p_i^* = \frac{1}{\bar{m}_{i-1} - \bar{m}_{i-2}} \left[\sum_{k=i-1}^{J-1} q_k(m_k - \bar{m}_{i-2}) - \sum_{k=i}^{J-1} p_{k+1}^*(\bar{m}_k - \bar{m}_{i-2})\right], \quad \text{for } i = 2, \cdots, J-1$$

$$p_1^* = \frac{1}{\bar{m}_0}\left[\sum_{k=0}^{J-1} q_k m_k - \sum_{k=1}^{N} p_{k+1}^* \bar{m}_k\right]$$

$$p_0^* = 1 - \sum_{i=1}^{J} p_i^*.$$

The proof of Theorem 4 is similar to that of Theorem 1 and, thus, is omitted.

The following two theorems prove that sum of optimum quantizations remains optimum for the aggregate traffic if each source is quantized individually. Moreover, the loss probability evaluated based on pseudosources is an upper bound of that evaluated based on real sources. Consider $K$ traffic sources $X_1, X_2, \cdots$ and $X_K$. Let $\mathbf{p_i}$ be a legal quantization of $X_i$ and let $\mathbf{p_i^*}$ be the optimum quantization. In the theorems $Z_i(\mathbf{p_i})$ denotes the traffic generated by the pseudosource obtained from quantization $\mathbf{p_i}$. Proofs of Theorems 5 and 6 are omitted because they are direct applications of variability ordering (see, for example, [9, Th. 2.2.3]).

*Theorem 5:* It holds that $E(\sum_{i=1}^K Z_i(\mathbf{p_i}) - c)^+ \ge E(\sum_{i=1}^K X_i - c)^+$ for all $c$.

*Theorem 6:* It holds that $E(\sum_{i=1}^K Z_i(\mathbf{p_i}^*) - c)^+ \le E(\sum_{i=1}^K Z_i(\mathbf{p_i}) - c)^+$ for all $c$ and $E(\sum_{i=1}^K Z_i(\mathbf{p_i}^*)) = E(\sum_{i=1}^K X_i)$.

## V. Numerical Examples

In the examples studied in this section we assume $C = 150$ Mb/s. The quantization levels are determined by a unit rate $u$. For example, if $u = 100$ kb/s, then $M_i = i \times 100$ kb/s, $0 \le i \le 1500$. Also, bandwidth is allocated in multiples of $u$. Three different values of $u$, i.e., 64, 100, and 200 kb/s, are used in our examples. Two types of traffic sources are considered. Type I and Type II sources are represented by $\{\mathbf{q_1}, \mathbf{m_1}, Y_1, 2\}$ and $\{\mathbf{q_2}, \mathbf{m_2}, Y_2, 3\}$, respectively, where $\mathbf{m_1} = [0\ 64]$ kb/s, $\mathbf{q_1} = [0.65\ 0.35]$, $\mathbf{m_2} = [1\ 1.53\ 4]$ Mb/s, and $\mathbf{q_2} = [0.4\ 0.5\ 0.1]$. The bufferless multiplexer system is considered and the desired cell loss probability (QoS) is restricted to be at most $10^{-9}$. The suboptimum quantization scheme stated in Theorem 4 is adopted in these examples.

In Tables I–IV we list the allocated bandwidth for various combinations of Type I and Type II sources. In these tables the values in the second column (i.e., $BW^*$) represent the minimum

TABLE I
Bandwidth Allocated When the Number of Type I Sources is 100 (Unit of $BW^*$ and $BW$ is Mb/s)

| Number of Type II Sources | | 64 Kbps | 100 Kbps | 200 Kbps |
|---|---|---|---|---|
| | $BW^*$ | $BW$ | $BW$ | $BW$ |
| 0 | 2.25 | 2.25 | 2.25 | 2.25 |
| 10 | 37.75 | 37.92 | 37.91 | 38.17 |
| 20 | 60.47 | 60.77 | 60.56 | 60.67 |
| 30 | 80.84 | 81.23 | 81.01 | 81.11 |
| 40 | 100.39 | 100.88 | 100.60 | 100.68 |

TABLE II
Bandwidth Allocated When the Number of Type I Sources is 500

| Number of Type II Sources | | 64 Kbps | 100 Kbps | 200 Kbps |
|---|---|---|---|---|
| | $BW^*$ | $BW$ | $BW$ | $BW$ |
| 0 | 11.24 | 11.24 | 11.24 | 11.24 |
| 10 | 46.88 | 47.08 | 47.25 | 48.10 |
| 20 | 69.36 | 69.66 | 69.64 | 70.21 |
| 30 | 89.78 | 90.17 | 90.08 | 90.56 |
| 40 | 109.33 | 109.82 | 109.66 | 110.08 |

TABLE III
Bandwidth Allocated When the Number of Type I Sources is 1000

| Number of Type II Sources | | 64 Kbps | 100 Kbps | 200 Kbps |
|---|---|---|---|---|
| | $BW^*$ | $BW$ | $BW$ | $BW$ |
| 0 | 22.48 | 22.48 | 22.48 | 22.48 |
| 10 | 58.27 | 58.46 | 58.85 | 60.38 |
| 20 | 80.58 | 80.87 | 81.06 | 82.17 |
| 30 | 100.99 | 101.39 | 101.47 | 102.40 |
| 40 | 120.55 | 121.04 | 121.02 | 121.84 |

TABLE IV
Bandwidth Allocated When the Number of Type I Sources is 2000

| Number of Type II Sources | | 64 Kbps | 100 Kbps | 200 Kbps |
|---|---|---|---|---|
| | $BW^*$ | $BW$ | $BW$ | $BW$ |
| 0 | 44.97 | 44.97 | 44.97 | 44.97 |
| 10 | 81.01 | 81.21 | 82.09 | 84.88 |
| 20 | 103.14 | 103.43 | 103.99 | 106.11 |
| 30 | 123.50 | 123.90 | 124.29 | 126.09 |

bandwidths required to meet the desired QoS, assuming the bandwidth is not quantized. For the quantized system, the minimum bandwidth required is denoted by $BW$. It can be verified that the percentage of loss, which is defined as $(BW - BW^*)/BW^*$, is small for all of the investigated cases. In other words, it is possible to choose a bigger $u$ to simplify the bandwidth allocation process with only a little sacrifice in utilization.

## VI. Conclusion

We have presented in this paper one method for characterizing (or quantizing) delay-sensitive traffic streams. Quantization does not alter characteristics of the real traffic, it only yields a pseudosource used by the network for resource allocation. Our proposed optimum quantization makes the network to reserve resource conservatively. However, our numerical examples show that, for a bufferless multiplexer, the bandwidth reserved based on pseudosources is only slightly greater than the minimum bandwidth required to meet the requested QoS. Taking into account buffering effect for a buffered system is an interesting research topic which can be further studied.

## References


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

[2] V. Anantharam and T. Konstantopoulos, "Burst reduction properties of leaky bucket flow control scheme in ATM networks," *IEEE Trans. Commun.*, vol. 42, pp. 3085–3089, 1994.

[3] J. Solé-Pareta and J. Domingo-Pascual, "Burstiness characterization of ATM cell streams," *Comput. Networks ISDN Syst.*, vol. 26, pp. 1351–1363, 1994.

[4] R. L. Cruz, "A calculus for network delay, part I: Network elements in isolation," *IEEE Trans. Inform. Theory*, vol. 37, pp. 114–131, 1991.

[5] H. Yong, W. Lee, and J. W. Mark, "ATM network traffic characterization using two types of on-off sources," in *Proc. IEEE INFOCOM*, **City, State/Country,** 1993, pp. 152–159.

[6] C. T. Lea and A. Alyatama, "Bandwidth quantization and states reduction in the broadband ISDN," *IEEE/ACM Trans. Networking*, vol. 3, pp. 352–360, June 1995.

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

[8] M. Shaked and J. G. Shanthikumar, *Stochastic Orders and Their Applications*. New York: Academic, 1994.

[9] D. Stoyan and D. J. Daley, *Comparison Methods for Queues and Other Stochastic Models*. New York: Wiley, 1983.



**Tsern-Huei Lee** (S'86–M'87) received the B.S. degree from National Taiwan University, Taipei, Taiwan, R.O.C., in 1981, the M.S. degree from University of California, Santa Barbara, in 1984, and the Ph.D. degree from University of Southern California, Los Angeles, in 1987, all in electrical engineering.

Since 1987 he has been a member of the Faculty of National Chiao Tung University, Hsinchu, Taiwan, R.O.C., where he is a Professor with the Department of Communication Engineering and a member of the Center for Telecommunications Research. His current research interests are in communication protocols, broad-band switch architectures, and traffic management.

Dr. Lee received an Outstanding Paper Award from the Institute of Chinese Engineers in 1991.



**Kuen-Chu Lai** received the B.S. degree in communication engineering and the Ph.D. degree in electronic engineering from National Chiao Tung University, Hsinchu, Taiwan, R.O.C., in 1989 and 1996, respectively.

He is currently a Consultant with Voice of Han Broadcasting Network, Taipei, Taiwan, R.O.C. His principal areas of interest include audio signal processing, traffic management, and traffic modeling.

Dr. Lai received the Broadcasting Engineering Contribution Prize from the Government Information Office in 1998.