

Performance analysis of statistical multiplexing of heterogeneous discrete-time Markovian arrival processes in an ATM network

Koohong Kang*, Cheeha Kim

Department of Computer Science and Engineering (NDS Lab), Pohang University of Science and Technology, San 31, Hyoja-Dong, Pohang 790-784, Korea

Received 14 June 1996; revised 3 May 1997; accepted 7 May 1997

Abstract

The cell loss probability and the mean cell delay are major performance metrics in analyzing a statistical multiplexer loaded with a superposition of independent and heterogeneous bursty sources. In this paper, we model each arrival process by a two-state discrete-time Markovian arrival process (D-MAP). We discuss that this traffic modeling is more realistic than the other ones in ATM networks. Then we model the superposition of r types of the two-state D-MAPs into a discrete-time batch Markovian arrival process (D-BMAP) with 2^r states. By using the steady-state analysis of the D-BMAP/D/1/K queueing model, we obtain the exact cell loss probabilities and the mean cell delays for each type of traffic in the statistical multiplexer. In particular, we derive the formulas concerned with these performance metrics under two buffer access strategies of the simultaneous cell arrivals at the same slot: (1) fair access, and (2) priority access. From some numerical examples, we show that the performance of each traffic at the statistical multiplexer may be severely affected by its own traffic characteristics and priority of buffer access, as well as the traffic characteristics of the others. © 1997 Elsevier Science B.V.

Keywords: Performance analysis; Discrete-time Markovian arrival process; ATM/B-ISDN

1. Introduction

Among the techniques proposed to implement broadband ISDN (B-ISDN), asynchronous transfer mode (ATM) is regarded as a desirable transfer mode which supports a wide variety of traffic for diverse service and performance requirements [1,2]. ATM is efficient to integrate bursty traffic and to achieve flexible bandwidth sharing because of its inherent capability of statistical multiplexing. But, before statistically multiplexing a new type of traffic with traffic existing in the network, it is necessary to estimate the characteristics of each traffic since its performance may be changed by the new type of traffic to be integrated in the network. The cells from the bursty traffic may be delayed and lost in the network because the allocated bandwidth is lower than its peak bit rate. Thus, the cell loss probability and the mean cell delay are major performance metrics in analyzing the ATM nodes (switch or multiplexer) [3–13].

The Poisson assumption used as an arrival process in the conventional packet-switched networks is no longer valid in ATM networks because the stream types of the packetized voice and video traffic as well as data traffic cause high correlation between adjacent cell arrivals [3,4,6,7]. Thus

an input process to a statistical multiplexer must be capable of representing such a correlation, which is crucial to understand the queueing behavior of the statistical multiplexer. For this reason, point processes [14] whose arrival rates vary randomly over time arise in many applications of interest, notably in communications modeling [4,7,8]. The versatile Markovian point process was introduced by Neuts in [15]. It is very rich class that contains many well known arrival point processes as special cases. This process has also been used extensively to model the input traffic for ATM networks because it models the time-varying arrival rate qualitatively and captures some of the important correlations between the interarrival times while still remaining analytically tractable [5–13].

In the literature, there are two main categories — continuous and discrete time models — for evaluating the performance of the ATM nodes. In continuous time models [3–9], it is assumed that cell arrivals can occur at any time instant. But in discrete time models [10–13], the cell arrivals and departures occur only at discrete time instants. Generally, the time axes of input and output lines of the ATM nodes are slotted with the slot size equal to the transmission time of a cell, and the slot boundaries on all input lines are synchronized [16–19]. Thus the actual cell arrivals and departures to and from the statistical multiplexer are

* Corresponding author.

embedded at slot boundaries and we believe that this discrete time model is more realistic in ATM networks.

In [10] and [20], the actual cell arrivals from a source to the ATM multiplexer were periodic or deterministic when the source is active. Even if the assumption of periodic cell arrival stream might be reasonable at the application levels, it will be no longer valid when its characteristics are distorted from the soft-multiplexings, which are done at the ATM adaptation layer (AAL) and the ATM layer in the end user system, and the shaper and policer at the customer premise.

In this paper, we consider a two-state discrete-time Markovian arrival process (D-MAP) [11] as the arrival process to an input port of the statistical multiplexer. A D-MAP which is equivalent to a Markov modulated Bernoulli process (MMBP) referred in to [13] is a doubly stochastic process whose rate varies according to the underlying Markov chain. The arrival process makes transitions among a finite number of states by a Markov process. The arrival rate at an arbitrary instant is determined by a state in which the process is. The D-MAP has an important advantage over other models because it captures the discrete-time nature and the bursty characteristics of the arrival cell stream. We consider the superposition of r types of the two-state D-MAPs which are synchronous for state transitions. This superposed bursty traffic is modeled using $m(=2')$ state discrete-time batch Markovian arrival process (D-MBAP [11,21]). New results for the m -state D-BMAP/D/1/K queueing model enable us to obtain the exact cell loss probabilities and the mean cell delays for each type of traffic in the statistical multiplexer loaded with a superposition of independent and heterogeneous bursty sources. In particular, we derive the formulas concerned with these performance metrics under two buffer access strategies of the simultaneous cell arrivals: (1) *fair access*, and (2) *priority access*.

The paper is organized as follows. In Section 2, we present the system and input traffic models. In Section 3, we analyze the system and obtain the cell loss probability and the mean cell delay for the superposition of input D-MAPs. In order to investigate the performance of each type of traffic separately, we derive the formulas for the cell loss probability and the mean cell delay for each of them under the fair access and the priority access strategies. In Section 4, we present some numerical results, and make concluding remarks in Section 5.

2. Model description

We model a statistical multiplexer by the queueing system as shown in Fig. 1. The statistical multiplexer transmits incoming cells from each of r input sources on the outgoing link. All incoming cells are stored in a common buffer, the capacity of which is K cells. When the buffer is full, arriving cells are blocked and lost, and those cells queued in the buffer are serviced on the first-come first-served (FCFS) basis. We consider that all input and output ports of the

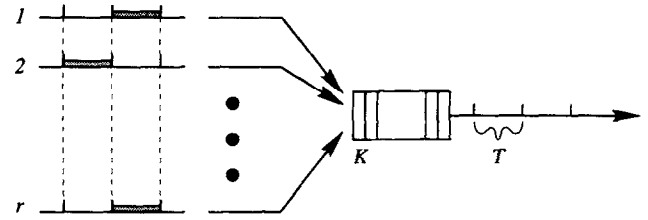


Fig. 1. Queueing model for a statistical multiplexer.

statistical multiplexer have the same transmission capacity, the time axis is slotted with the slot size equal to the transmission time T of a cell, and that the slot boundaries on all input ports are synchronized (i.e., the operation of the statistical multiplexer is synchronous). Hence all new cells are assumed to arrive at the beginning of the slot and a cell stored in the buffer departs at the end of the slot. Each input traffic to the statistical multiplexer is modeled by a two-state D-MAP. As shown in Fig. 2, the input traffic has two states, S_1 and S_2 . When the input traffic is in S_1 (or S_2) state, one cell is generated with a constant probability p_1 (or p_2) in each time slot. Suppose that the input source is in S_1 (or S_2) state in the time slot s . Then, in the next time slot $s+1$, it will change to S_2 (or S_1) state with probability α_1 (or α_2), or it will remain in S_1 (or S_2) state with probability $1-\alpha_1$ (or $1-\alpha_2$).

In this paper, we suppose that the r types of arrival processes, one per input port, are independent and heterogeneous from each other. Then, the two-state D-MAP of arrival process j , $j=1,2,\dots,r$, is characterized by a transition probability matrix M_j of the underlying Markov process and a cell generation probability matrix P_j in each state and slot, which have the form

$$M_j = \begin{bmatrix} 1 - \alpha_{1,j} & \alpha_{1,j} \\ \alpha_{2,j} & 1 - \alpha_{2,j} \end{bmatrix}, \quad P_j = \begin{bmatrix} p_{1,j} & 0 \\ 0 & p_{2,j} \end{bmatrix},$$

$$j = 1, 2, \dots, r. \quad (1)$$

The superposition of r types of these two-state D-MAPs parameterized by (M_j, P_j) , $1 \leq j \leq r$, which are independent and synchronous with each other at the time slot instant, yields a D-BMAP with transition matrix D having $m(=2')$ states as follows

$$D = M_1 \otimes M_2 \otimes \dots \otimes M_r, \quad (2)$$

where \otimes is the Kronecker product. The Kronecker product [22] of an $(n_1 \times m_1)$ matrix A and an $(n_2 \times m_2)$ matrix B is the $(n_1 n_2 \times m_1 m_2)$ matrix $A \otimes B$ with elements $A_{ij} B_{kl}$.

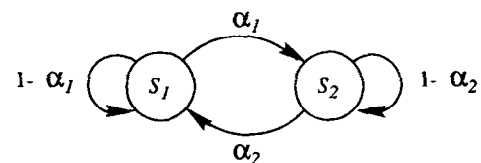


Fig. 2. Two-state discrete-time Markovian arrival process.

For the D-BMAP, the steady-state probability vector θ of the Markov process is given by solving the following equations:

$$\theta D = \theta, \quad (3)$$

$$\theta e = 1,$$

where e is a column vector of all ones.

Suppose that at time slot s the D-BMAP is in some state i , $1 \leq i \leq m$. At the next time slot $s + 1$, a transition to another or possibly the same state takes place and a batch arrival may or may not occur. With probability $(d_0)_{i,j}$, $1 \leq j \leq m$, there is a transition to state j without an arrival, and with probability $(d_n)_{i,j}$, $1 \leq j \leq m$, there is a transition to state j with a batch arrival of size n , $1 \leq n \leq r$. We have that

$$\sum_{n=0}^r \sum_{j=1}^m (d_n)_{i,j} = 1. \quad (4)$$

Clearly the matrix D_0 with elements $(d_0)_{i,j}$ governs transitions that correspond to no arrivals, while the matrices D_n with elements $(d_n)_{i,j}$, $1 \leq n \leq r$, govern transitions that correspond to arrivals of batches of size n . Hence the $(m \times m)$ matrix, D_n , $n = 0, 1, \dots, r$, is a matrix whose (i, j) th element is the probability that during a time slot there are n arrivals and that at the end the phase is j , given that at the beginning of the slot the phase was i .

Let $c_k(i)$ denote the probability of having k arrivals during a time slot when D-BMAP is in i state, then we can obtain $c_k(i)$, $k = 0, 1, \dots, r$, of which the computational complexity is $O(r^2)$ (see Appendix A). Hence D_n is given by

$$D_n = C_n D, \quad (5)$$

where C_n is the following diagonal matrix

$$C_n = \text{Diag}(c_n(1), \dots, c_n(i), \dots, c_n(m)). \quad (6)$$

The effective arrival rate λ^* of the D-BMAP, which is a superposition of r types of the two-state D-MAPs, is given by

$$\lambda^* = \theta \sum_{n=0}^r n D_n e. \quad (7)$$

and the arrival rate from the two-state D-MAP j , $1 \leq j \leq r$, input traffic is represented by

$$\begin{aligned} P(j) &= \overbrace{0 \oplus \dots \oplus 0 \oplus P_j \oplus 0 \oplus \dots \oplus 0}^{r \text{ streams}} \\ &= I_2 \otimes \dots \otimes I_2 \otimes P_j \otimes I_2 \otimes \dots \otimes I_2, \end{aligned}$$

where \oplus is the Kronecker sum [22] and I_2 is the (2×2) identity matrix. The Kronecker sum \oplus is defined as $A \oplus B = (A \otimes I_B) + (I_A \otimes B)$, where I_A and I_B are the identity matrices of the same order as the matrices A and B respectively. The effective arrival rate from the two-state D-MAP j input traffic is also given by

$$\lambda_j^* = \theta P(j) e, \quad j = 1, 2, \dots, r. \quad (9)$$

3. Performance analysis

As shown in Fig. 1, the cells may arrive simultaneously at the same slot on input lines of the statistical multiplexer. Since we focus on the performance analysis of individual traffic types, the buffer access strategies for those cells are crucial as will be shown in the numerical results, even if the mean performance of the superposed traffic is not affected by the access strategies. Let us define two access strategies for such cells as follows:

- fair access strategy — the cells arriving at the same slot join the finite capacity buffer with equal probability, hence they are blocked and lost randomly as much as the excess cells are overflowed owing to buffer-full;
- priority access strategy — the cells arriving at the same slot join the finite capacity buffer with pre-determined priorities.

3.1. Queue length distribution embedded at slot ending-time instant

Consider the statistical multiplexer whose input consists of the superposition of the r types of the two-state D-MAPs as described in the previous section. We consider this multiplexer at discrete time epochs $0, T, 2T, \dots$. Hence the multiplexer is modeled as a D-BMAP/D/1/K queue, where the service time is equal to one time slot. Let L_n and J_n be, respectively, the number of cells queued in the buffer excluding a cell which has just been removed for transmission and the state of the input process at time instant nT . We assume that an arriving cell departs at the end of its arriving time slot if the buffer is empty. Then, the process $\{(L_n, J_n; n \geq 0)\}$ forms a finite-state Markov chain with state space $\{0, 1, \dots, K-1\} \times \{1, \dots, m\}$, and the transition probability matrix Q^K of the Markov chain is given by

$$Q^K = \begin{bmatrix} D_0 + D_1 & D_2 & \dots & D_{K-1} & \sum_{k=K}^r D_k \\ D_0 & D_1 & \dots & D_{K-2} & \sum_{k=K-1}^r D_k \\ 0 & D_0 & \dots & D_{K-3} & \sum_{k=K-2}^r D_k \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & D_0 & \sum_{k=1}^r D_k \end{bmatrix}. \quad (10)$$

For the Markov chain with this transition probability matrix Q^K , define the limiting probability distribution $x_{k,i}$ as

$$x_{k,i} = \lim_{n \rightarrow \infty} \text{Prob}\{L_n = k, J_n = i\}, \quad (11)$$

and let the vector $\mathbf{x} = (x_0, x_1, \dots, x_{K-1})$ be the steady-state probability vector with $\mathbf{x}_k = (x_{k,1}, x_{k,2}, \dots, x_{k,m})$, $0 \leq k \leq K-1$. Then the steady-state probability vector \mathbf{x} is obtained

from the equations

$$\mathbf{x}Q^K = \mathbf{x}, \quad \mathbf{x}\mathbf{e} = 1. \quad (12)$$

3.2. Performance measures under the fair access strategy

The following quantities are easily calculated by using the queue length distribution obtained previously. The mean number of cells lost due to buffer overflow is given by

$$L = \sum_{l=K+1}^r (l-K)x_0 D_l e + \sum_{k=1}^{K-1} \sum_{l=1}^r [l-K+k]^+ x_k D_l e, \quad (13)$$

where $[x]^+ = \max(x, 0)$. The cell loss probability P_b for an arbitrary cell of the superposition of input B-MAPs is obtained by

$$P_b = \frac{L}{\lambda^*}. \quad (14)$$

In the fair access strategy, since cells arriving at the same slot join the finite capacity buffer with equal probability, they are blocked and lost randomly as much as the excess cells are overflowed owing to buffer-full. Hence the mean number of cells lost arriving from the two-state D-MAP j is given by

$$L^j = \sum_{l=K}^{r-1} \frac{l-K+1}{l+1} x_0 P(j) D_l^j e + \sum_{k=1}^{K-1} \sum_{l=1}^{r-1} \frac{[l-K+k+1]^+}{l+1} x_k P(j) D_l^j e, \quad (15)$$

$j = 1, 2, \dots, r,$

where the matrices D_n^j govern transitions that correspond to arrivals of batches of size n excluding the arrival from the two-state D-MAP j (see Appendix B). The cell loss probability P_b^j for cells arriving from the type j traffic, $1 \leq j \leq r$, is then given by

$$P_b^j = \frac{L^j}{\lambda_j^*}. \quad (16)$$

For the mean cell delay W for cells of the superposition of input D-MAPs, we consider $n, n = 1, 2, \dots, r$, cells arriving at time kT and tag them. The tagged cells may find the system in one of the following states.

- There is no cell queued in the buffer. In this case, if $n \leq K$, the total waiting time of each arrival of cells is $n(n+1)T/2$. If $n > K$, $n-K$ cells are blocked. The waiting time of a blocked cell can be regarded as zero. Hence if $n > K$, the total waiting time of each arrival of non-blocked cells is $K(K+1)T/2$.
- There are $m, 1 \leq m \leq K-1$, cells queued in the buffer. In this case, the tagged cells may have to wait until these cells are served. Hence if $n \leq K-m$, the total waiting time of each arrival of cells is $\{nm + [n(n+1)/2]\}T$. If

$n > K-m$, $n-K+m$ cells are blocked. Hence if $n > K-m$, the total waiting time of each arrival of non-blocked cells is $\{(K-m)m + [(K-m)(K-m+1)/2]\}T$.

Finally, we can obtain the mean cell delay W for an arbitrary cell which is not blocked as follows

$$W = \frac{T_0 + \sum_{m=1}^{K-1} T_m}{(1-P_b)\theta \sum_{n=0}^r n D_n e}, \quad (17)$$

where $T_m, m = 0, 1, \dots, K-1$, denotes the total waiting time of each arrival of cells which is non-blocked and finds that the queue length is m . Hence T_m s are given by

$$T_0 = \sum_{n=1}^K \frac{n(n+1)T}{2} x_0 D_n e + \sum_{n=K+1}^r \frac{K(K+1)T}{2} x_0 D_n e, \quad (18)$$

$$\begin{aligned} \sum_{m=1}^{K-1} T_m &= \sum_{m=1}^{K-1} \left[\sum_{n=1}^{K-m} \left\{ nm + \frac{n(n+1)}{2} \right\} T x_m D_n e \right. \\ &\quad \left. + \sum_{n=K-m+1}^r \left\{ (K-m)m + \frac{(K-m)(K-m+1)}{2} \right\} \right. \\ &\quad \left. \times T x_m D_n e \right]. \end{aligned} \quad (19)$$

Similarly, for the mean cell delay W^j for cells arriving from the type j traffic, we consider an arbitrary cell arriving from the two-state D-MAP j traffic at time kT and tag it. Because all new cells from the two-state D-MAPs are assumed to arrive at the beginning of the time slot simultaneously, the mean waiting time of the tagged cell is dependent on the arrivals of batches of size n excluding the arrival from the two-state D-MAP j as well as the queue length distribution that the tagged arrival of cell finds. The tagged cell may find the system in one of the following states.

- There is no cell queued in the buffer. In this case, if $n \leq K-1$, the tagged cell will be served at an arbitrary time slot between T and $(n+1)T$. Hence the mean waiting time of the tagged cell will be $(n+2)T/2$. If $n \geq K$, $n+1-K$ cells are blocked and the tagged cell joins the buffer with probability $1 - [(n+1-K)/(n+1)]$. Hence if $n \geq K$, the waiting time of a blocked tagged cell can be regarded as zero, and the mean waiting time of the non-blocked tagged cell is $(K+1)T/2$.
- There are $m, 1 \leq m \leq K-1$, cells queued in the buffer. In this case, the tagged cell may have to wait until these cells are served, hence if $n \leq K-m-1$, the mean waiting time of the tagged cell will be $[m + (n+2)/2]T$. If $n \geq K-m$, $n+1-K-m$ cells are blocked and the tagged cell joins the buffer with probability $1 - [(n+1-K+m)/(n+1)]$. Hence the mean waiting time of the non-blocked tagged cell is $[m + (K-m+1)/2]T$.

Finally, we can obtain the mean cell delay W^j for cells arriving from D-MAP j which are not blocked as follows

$$W^j = \frac{T_0^j + \sum_{m=1}^{K-1} T_m^j}{(1 - P_b^j)\theta P(j)e}, \quad j = 1, 2, \dots, r, \quad (20)$$

where T_m^j s are given by

$$T_0^j = \sum_{n=0}^{K-1} \frac{(n+2)T}{2} x_0 P(j) D_n^j e + \sum_{n=K}^{r-1} \frac{K(K+1)T}{2(n+1)} x_0 P(j) D_n^j e, \quad (21)$$

$$\begin{aligned} \sum_{m=1}^{K-1} T_m^j = & \sum_{m=1}^{K-1} \left[\sum_{n=0}^{K-m-1} \left(m + \frac{n+2}{2} \right) T x_m P(j) D_n^j e \right. \\ & \left. + \sum_{n=K-m}^{r-1} \frac{(K-m)(K+m+1)}{2(n+1)} T x_m P(j) D_n^j e \right]. \end{aligned} \quad (22)$$

3.3. Performance measures under the priority access strategy

Now, in order to analyze the performance of the priority access strategy, we assume, *w.l.o.g.*, that the priority of type j is higher than that of type i to access the finite capacity buffer when $1 \leq j < i \leq r$. Hence the mean number of cells lost arriving from the D-MAP j is given by

$$L_h^j = \sum_{k=0}^{K-1} \sum_{s=K-k}^{j-1} x_k P(j) G_s(1, j-1), \quad j = 1, 2, \dots, r, \quad (23)$$

where $G_s(1, j-1)$ is the following $(m \times m)$ diagonal matrix

$$G_s(1, j-1) = \text{Diag}[G_{s(1)}(1, j-1), G_{s(2)}(1, j-1), \dots, G_{s(m)}(1, j-1)], \quad (24)$$

where $G_{s(i)}(x, y)$ as the probability of generation of s cells during a time slot between two-state D-MAP x and y , when D-BMAP is i state (see Appendix A). We note that if $j = 1$, then $L_h^1 = 0$. The cell loss probability P_b^j for cells arriving from the type j traffic considering the buffer access priority is then given by

$$P_b^j = \frac{L_h^j}{\lambda_j^*}, \quad j = 1, 2, \dots, r. \quad (25)$$

The mean waiting time W_h^j for cells arriving from D-MAP j which are not blocked is as follows

$$W_h^j = \frac{T_h^j}{(1 - P_b^j)\theta P(j)e} \quad (26)$$

where T_h^j is given by

$$T_h^j = \sum_{m=0}^{K-1} \sum_{n=0}^{j-1} (n+1+m) T x_m P(j) G_n(1, j-1). \quad (27)$$

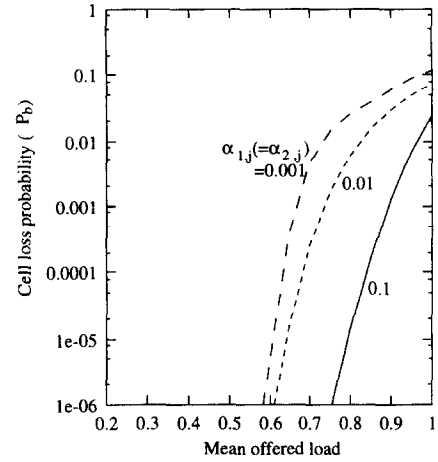


Fig. 3. Effects of the arrival state duration parameters $\alpha_{1,j}$ and $\alpha_{2,j}$ on the cell loss probability (P_b) for three homogeneous two-state D-MAPs with $K = 25$ and $p_{1,j}/p_{2,j} = 3$.

4. Numerical results and discussion

In this section, some numerical results are presented to illustrate the effects of each traffic's characteristic and the two buffer access strategies on the performance of statistical multiplexing.

We investigate the performance of a statistical multiplexer with three two-state D-MAPs. These D-MAPs are parameterized by (M_1, P_1) , (M_2, P_2) and (M_3, P_3) as shown in Eq. (1), respectively. In Figs. 3 and 4, we show the cell-loss probability P_b and the mean cell delay W for an arbitrary cell when these D-MAPs are homogeneous and aggregated, for which the ratio of cell generation probabilities in each state $p_{1,j}/p_{2,j}$ is kept constant at 3 with $\alpha_{1,j}$ and $\alpha_{2,j}$ as parameters. The effective arrival rate λ_j^* of the two-state D-MAP j is given by

$$\lambda_j^* = \frac{\alpha_{1,j} p_{2,j} + \alpha_{2,j} p_{1,j}}{\alpha_{1,j} + \alpha_{2,j}}, \quad j = 1, 2, 3, \quad (28)$$

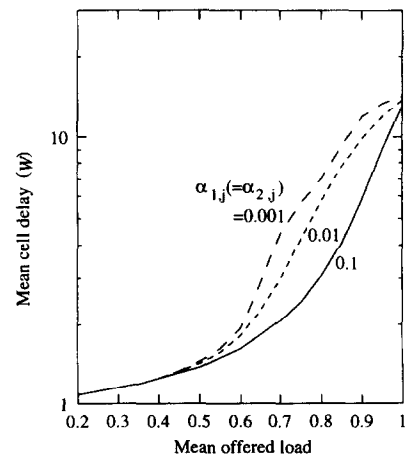


Fig. 4. Effects of the arrival state duration parameters $\alpha_{1,j}$ and $\alpha_{2,j}$ on the mean cell delay (W) for three homogeneous two-state D-MAPs with and $p_{1,j}/p_{2,j} = 3$.

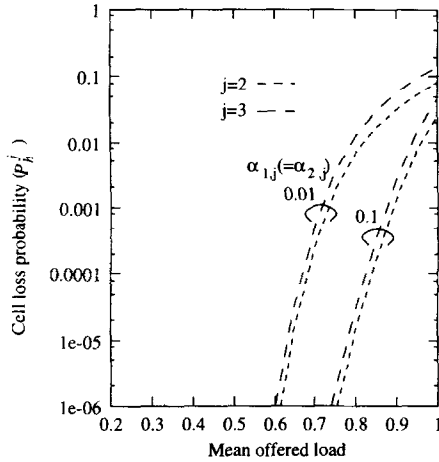


Fig. 5. A comparison of the cell loss probabilities (P_h^1, P_h^2, P_h^3) for three homogeneous two-state D-MAPs with $K = 25$ and $p_{1,j}/p_{2,j} = 3$ in case of the priority access strategy, where $P_h^1 = 0$.

and the squared coefficient of variation of the interarrival time, C_j^2 , which can be used as the basis for determining the degree of bursty nature of the cell arrival process, is given by (see [13])

$$C_j^2 = \frac{2\lambda_j^* \left[(\alpha_{1,j} + \alpha_{2,j})^2 + (\alpha_{1,j}p_{1,j} + \alpha_{2,j}p_{2,j})(1 - \alpha_{1,j} - \alpha_{2,j}) \right]}{(\alpha_{1,j} + \alpha_{2,j}) [\alpha_{2,j}p_{1,j} + \alpha_{1,j}p_{2,j} + p_{1,j}p_{2,j}(1 - \alpha_{1,j} - \alpha_{2,j})]} - \lambda_j^* - 1, \quad j = 1, 2, 3. \quad (29)$$

In this example and the following ones, we set the slot duration T equal to one, so that the unit of time for the mean cell delay is the slot duration. From these figures, as could be expected, the cell loss probability and the mean cell delay increase as the mean sojourn times in each arrival state of the two-state D-MAP, i.e., the inverses of $\alpha_{1,j}$ and $\alpha_{2,j}$, increase. This means that an increase in the burstiness of a cell stream results in a worse performance at the multiplexer.

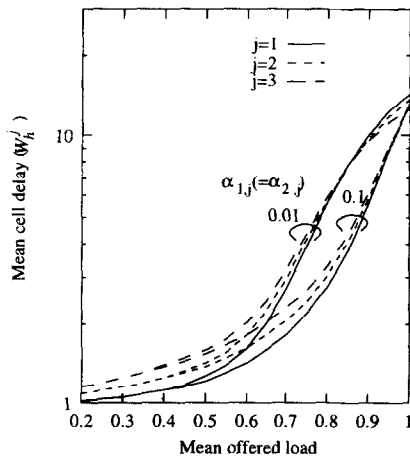


Fig. 6. A comparison of the mean cell delays (W_h^1, W_h^2, W_h^3) for three homogeneous two-state D-MAPs with $K = 25$ and $p_{1,j}/p_{2,j} = 3$ in case of the priority access strategy.

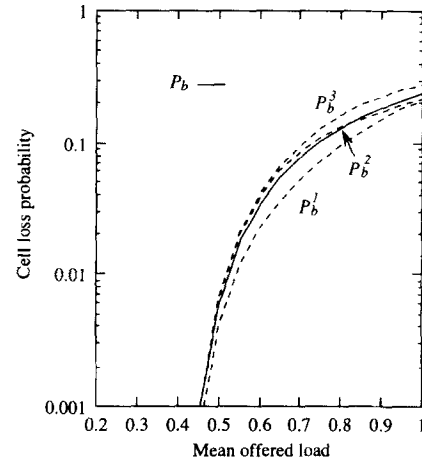


Fig. 7. A comparison of the cell loss probabilities (P_h, P_b^1, P_b^2, P_b^3) for three heterogeneous D-MAPs with parameters $(\alpha_{1,1}, \alpha_{2,1}) = (0.2, 0.1)$, $(\alpha_{1,2}, \alpha_{2,2}) = (0.02, 0.01)$, $(\alpha_{1,3}, \alpha_{2,3}) = (0.002, 0.001)$, $p_{2,1} = p_{2,2} = p_{2,3} = 0$, and $K = 25$ in case of the fair access strategy.

In Figs. 5 and 6, we also show the cell loss probability P_h^j and the mean cell delay W_h^j for the two-state D-MAP $j, j = 1, 2, 3$, at a statistical multiplexer employing the priority access strategy. We assume that type j 's priority is higher than type i 's to access the finite capacity buffer when $1 \leq j < i \leq 3$. As could be also expected, the priority access strategy discriminates the performances of each D-MAP based on its priority even though the input sources are homogeneous, while the numerical results of the fair access strategy are the same as shown in Figs. 3 and 4. In particular, P_h^1 results in zero because the cells from type 1 always find the buffer occupancy distribution such that $\text{Prob}[x < K] = 1$, where K is buffer size.

In Figs. 7 and 8, we investigate the performance of a statistical multiplexer employing the fair access strategy with the three heterogeneous types of traffic; we have $(\alpha_{1,1}, \alpha_{2,1}) = (0.2, 0.1)$, $(\alpha_{1,2}, \alpha_{2,2}) = (0.02, 0.01)$, $(\alpha_{1,3}, \alpha_{2,3}) =$

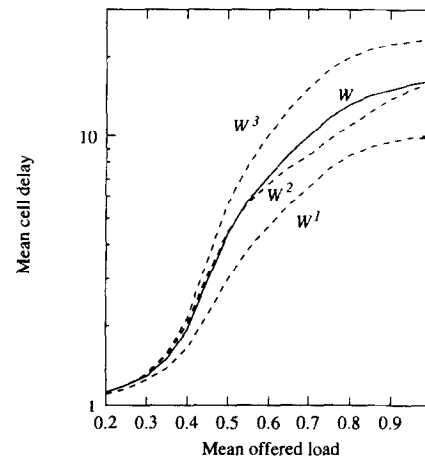


Fig. 8. A comparison of the cell loss probabilities (W, W^1, W^2, W^3) for three heterogeneous D-MAPs with parameters $(\alpha_{1,1}, \alpha_{2,1}) = (0.2, 0.1)$, $(\alpha_{1,2}, \alpha_{2,2}) = (0.02, 0.01)$, $(\alpha_{1,3}, \alpha_{2,3}) = (0.002, 0.001)$, $p_{2,1} = p_{2,2} = p_{2,3} = 0$, and $K = 25$ in case of the fair access strategy.

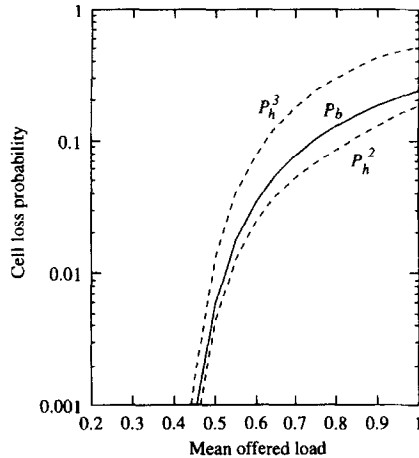


Fig. 9. A comparison of the cell loss probabilities (P_b, P_h^1, P_h^2, P_h^3) for three heterogeneous D-MAPs with parameters $(\alpha_{1,1}, \alpha_{2,1}) = (0.2, 0.1)$, $(\alpha_{1,2}, \alpha_{2,2}) = (0.02, 0.01)$, $(\alpha_{1,3}, \alpha_{2,3}) = (0.002, 0.001)$, $p_{2,1} = p_{2,2} = p_{2,3} = 0$, and $K = 25$ in case of the priority access strategy, where $P_h^1 = 0$.

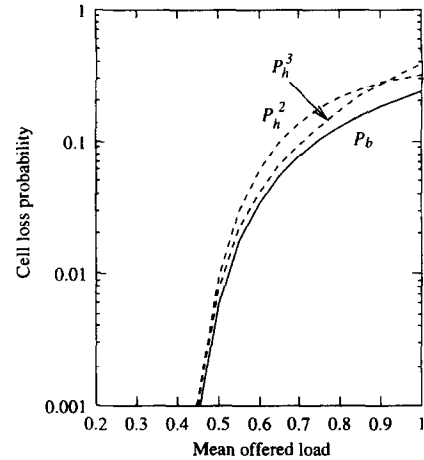


Fig. 11. A comparison of the cell loss probabilities (P_b, P_h^1, P_h^2, P_h^3) for three heterogeneous D-MAPs with parameters $(\alpha_{1,1}, \alpha_{2,1}) = (0.002, 0.001)$, $(\alpha_{1,2}, \alpha_{2,2}) = (0.02, 0.01)$, $(\alpha_{1,3}, \alpha_{2,3}) = (0.2, 0.1)$, $p_{2,1} = p_{2,2} = p_{2,3} = 0$, and $K = 25$ in case of the priority access strategy, where $P_h^1 = 0$.

$= (0.002, 0.001)$ and vary the arrival probabilities $p_{1,1}$, $p_{1,2}$ and $p_{1,3}$ while setting $p_{2,1}$, $p_{2,2}$ and $p_{2,3}$ equal to zero. From these figures, we can see that the cell loss probability and the mean cell delay are different from each type of traffic in the heterogeneous traffic case, but there are no sharp differences of loss performance based on the traffic characteristics of each type as we might expect in Fig. 3. This means that the more bursty stream influences severely the overall performance of the multiplexer and the less bursty streams are penalized. Moreover, this is because the behavior of cell loss at the multiplexer is bursty as much as input streams, such that the influence of the more bursty stream is dominant.

In Figs. 9–12, we investigate the performance of a statistical multiplexer employing the priority access strategy with three heterogeneous types of traffic. In order to investigate the effect of this buffer access strategy on the performance

in detail, we have $(\alpha_{1,1}, \alpha_{2,1}) = (0.2, 0.1)$, $(\alpha_{1,2}, \alpha_{2,2}) = (0.02, 0.01)$, $(\alpha_{1,3}, \alpha_{2,3}) = (0.002, 0.001)$ in Figs. 9 and 10, and $(\alpha_{1,1}, \alpha_{2,1}) = (0.002, 0.001)$, $(\alpha_{1,2}, \alpha_{2,2}) = (0.02, 0.01)$, $(\alpha_{1,3}, \alpha_{2,3}) = (0.2, 0.1)$ in Figs. 11 and 12, in which the characteristics of each type of traffic are fully reversed. We note that P_h^1 equals to zero in Figs. 9, and 11. We also note that the mean delays of type 3 in Figs. 10, and 12 do not increase monotonically for increasing mean offered load due to the loss of cells, and that the more bursty traffic gets high priority to buffer access, the more discrimination of the delay performance occurs. From these figures, we conclude that in the heterogeneous traffic case, the performance of each traffic at the statistical multiplexer is different, depending strongly on its traffic characteristic and the priority of buffer access, as well as on the traffic characteristics of the others.

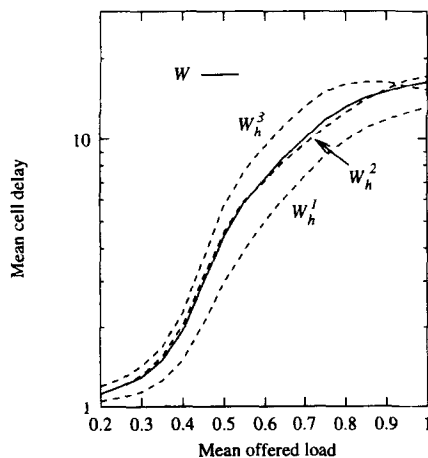


Fig. 10. A comparison of the mean cell delays (W, W_h^1, W_h^2, W_h^3) for three heterogeneous D-MAPs with parameters $(\alpha_{1,1}, \alpha_{2,1}) = (0.2, 0.1)$, $(\alpha_{1,2}, \alpha_{2,2}) = (0.02, 0.01)$, $(\alpha_{1,3}, \alpha_{2,3}) = (0.002, 0.001)$, $p_{2,1} = p_{2,2} = p_{2,3} = 0$, and $K = 25$ in case of the priority access strategy.

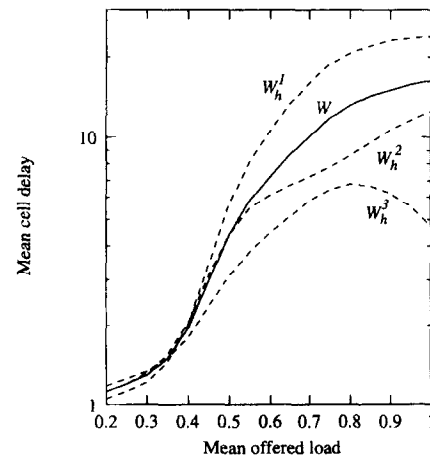


Fig. 12. A comparison of the mean cell delays (W, W_h^1, W_h^2, W_h^3) for three heterogeneous D-MAPs with parameters $(\alpha_{1,1}, \alpha_{2,1}) = (0.002, 0.001)$, $(\alpha_{1,2}, \alpha_{2,2}) = (0.02, 0.01)$, $(\alpha_{1,3}, \alpha_{2,3}) = (0.2, 0.1)$, $p_{2,1} = p_{2,2} = p_{2,3} = 0$, and $K = 25$ in case of the priority access strategy.

5. Conclusion

We have analyzed the performance of a statistical multiplexer for heterogeneous types of bursty traffic in an ATM network. We first modeled each bursty traffic by a two-state D-MAP, and discussed that this two-state D-MAP modeling is more realistic than the other existing traffic models in ATM networks. The different burstiness of each traffic was represented by changing the parameters of the two-state D-MAP. The superposition of r types of these two-state D-MAPs was modeled by a D-BMAP with $m(=2^r)$ states. With these traffic and system modeling, we obtained the exact formulas for the cell loss probability and the mean cell delay for the superposed heterogeneous bursty traffic. We also derived formulas for these performance metrics of each type of traffic separately under two buffer access strategies of the simultaneous cell arrivals: (1) fair access, and

yields a D-BMAP having $m(=2^r)$ states. We assume that the probability of cell generation of the two-state D-MAP j is $p_{j(i)}$ when the D-BMAP is in the i , $1 \leq i \leq m$, state, where $j(i) = 1$ (or 2) from Eq. (1). Let $c_k(i)$ denote the probability of having k arrivals during a time slot when the D-BMAP is i state. We assume, *w.l.o.g.*, that these two-state D-MAPs indexed by j , $j = 1, 2, \dots, r$, are ordered in increasing order, and define $G_{s(i)}(x, y)$ as the probability of s cells generation during a time slot between the D-MAP x and y , $1 \leq x \leq y \leq r$, when the D-BMAP is i state. Then we can obtain $G_{s(i)}(x, y)$, $0 \leq s \leq y - x + 1$, by the following recursive formulas:

$$s = 0; \quad G_{0(i)}(x, y) = \prod_{l=x}^y (1 - p_{l(i), l}) \quad (A1)$$

$$s = 1; \quad G_{1(i)}(x, y) = \begin{cases} p_{x(i), x} & \text{if } y = x, \\ p_{x(i), x} \cdot G_{0(i)}(x + 1, y) + \sum_{l=x+1}^{y-1} G_{0(i)}(x, l - 1) \cdot p_{l(i), l} \cdot G_{0(i)}(l + 1, y) + G_{0(i)}(x, y - 1) \cdot p_{y(i), y} & \text{if } y - x \geq 1, \end{cases} \quad (A2)$$

$$2 \leq s \leq r; \quad G_{s(i)}(x, y) = \begin{cases} \prod_{l=x}^y p_{l(i), l} & \text{if } y - x + 1 = s, \\ p_{x(i), x} \cdot G_{s-1(i)}(x + 1, y) + \sum_{l=x+1}^{y-s+1} G_{0(i)}(x, l - 1) \cdot p_{l(i), l} \cdot G_{s-1(i)}(l + 1, y) & \text{if } y - x + 1 \geq s + 1, \\ 0 & \text{otherwise.} \end{cases} \quad (A3)$$

(2) priority access.

On the basis of some numerical examples, we showed the effects of the characteristics of each traffic on the performance of the statistical multiplexer. In particular, we showed the effects of the two different buffer access strategies of the simultaneous arrival of cells on these performance metrics. We conclude that the performance of each traffic at the statistical multiplexer may be severely affected by its own traffic characteristics and priority of buffer access, as well as on the characteristics of the other types of traffic. We also note that both access strategies should be determined carefully at the hardware implementation of the statistical multiplexer for their purpose such as priority control or fair access control.

Appendix A

The superposition of r types of the two-state D-MAPs parameterized by (D_j, P_j) , $1 \leq j \leq r$, which are independent and synchronous with each other at the time slot instants,

Finally, we can obtain $c_k(i) = G_{k(i)}(1, r)$ for $k = 0, 1, \dots, r$ and $i = 1, 2, \dots, m$. We can notice that $G_{k(i)}(1, r)$, $1 \leq k \leq r$, is given recursively by $G_{k-1(i)}(t, r)$, $2 \leq t \leq r - k + 2$, and $G_{0(i)}(1, t)$, $1 \leq t \leq r - k$. Since the computational complexity for calculating $G_{k-1(i)}(t, r)$, $2 \leq t \leq r - k + 2$, by Eqs. (A1), (A2) and (A3) is $O(r)$, the computation complexity for obtaining $c_k(i)$, $k = 0, 1, \dots, r$, is $O(r^2)$.

Appendix B

Let $c_k^j(i)$, $j = 1, 2, \dots, r$, denote the probability of having k arrivals during a time slot when the D-BMAP is state i excluding the arrival from the two-state D-MAP j . From Eqs. (A1), (A2) and (A3) discussed in Appendix A, $c_k^1(i)$ and $c_k^r(i)$ at the boundaries ($j = 1, r$) are obtained as follows

$$c_k^1(i) = G_{k(i)}(2, r), \quad (B1)$$

and

$$c_k^r(i) = G_{k(i)}(1, r - 1). \quad (B2)$$

If $r > 2$ and $j \neq 1$ (or r), then $c_k^j(i)$ are obtained as follows

$$c_k^j(i) = \sum_{t=0}^k G_{t(i)}(1, j-1) G_{k-t(i)}(j+1, r), \quad j=2, 3, \dots, r-1. \quad (B3)$$

Similar to D_n in Eq. (5), matrices D_n^j govern transitions that correspond to arrivals of batches of size n excluding the arrival from the two-state D-MAP j . Hence D_n^j , $n = 0, 1, \dots, r-1$ and $j = 1, \dots, r$, is a matrix whose (i, k) th element is the probability that during a time slot there are n arrivals except from the traffic type j and that at the end of the phase is k , given that at the beginning of the slot the phase was i . Finally, D_n^j is given by

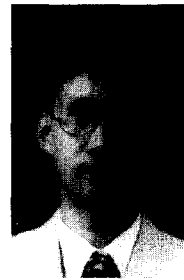
$$D_n^j = C_n^j D \quad (B4)$$

where C_n^j is the following diagonal matrix

$$C_n^j = \text{Diag}(c_n^j(1), \dots, c_n^j(i), \dots, c_n^j(m)). \quad (B5)$$

References

- [1] J.J. Bae, T. Suda, Survey of traffic control schemes and protocols in ATM networks, *Proc. IEEE* 79 (2) (1991) 170–189.
- [2] D.E. McDysan, D.L. Spohn, *ATM Theory and Application*, McGraw-Hill Inc., 1994.
- [3] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson, J.D. Robbins, Performance models of statistical multiplexing in packet video communications, *IEEE Trans. Commun.* 36 (7) (1988) 834–844.
- [4] K. Sriram, W. Whitt, Characterizing superposition arrival process in packet multiplexers for voice and data, *IEEE J. Selected Areas Commun.* 4 (6) (1986) 833–846.
- [5] A. Baiocchi, N. Bléfari-Melazzi, Steady-state analysis of the MMPP/G/1/K queue, *IEEE Trans. Commun.* 41 (9) (1993) 531–534.
- [6] A. Baiocchi, N. Bléfari-Melazzi, M. Listanti, A. Roveri, R. Winkler, Loss performance analysis of an ATM multiplexer loaded with high-speed ON-OFF sources, *IEEE J. Selected Areas Commun.* 9 (3) (1991) 388–393.
- [7] D. Heffes, D.M. Lucantoni, A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer performance, *IEEE J. Selected Areas Commun.* 4 (6) (1986) 856–868.
- [8] W. Fischer, K. Meier-Hellstern, The Markov-modulated Poisson process (MMPP) cookbook, *Performance Eval.* 18 (1992) 149–172.
- [9] Y.H. Kim, C.Y. Un, Performance analysis of statistical multiplexing for heterogeneous bursty traffic in an ATM network, *IEEE Trans. Commun.* 42 (2/3/4) (1994) 745–753.
- [10] T. Yang, D.H.K. Tsang, A novel approach to estimating the cell loss probability of an ATM multiplexer loaded with homogeneous bursty sources, in: *Proceedings of GLOBECOM'92*, 1992, pp. 511–517.
- [11] C. Blondia, O. Casals, Performance analysis of statistical multiplexing of VBR sources, in: *Proceedings of INFOCOM'92*, 1992, pp. 6C.2.1 to 6C.2.11.
- [12] S.S. Wang, J.A. Silvester, A discrete-time performance model for integrated service ATM multiplexers, in: *Proceedings of GLOBECOM'93*, 1993, pp. 757–761.
- [13] D. Park, H.G. Perros, m-MMBP Characterization of the departure process of an m-MMBP/Geo/1/K queue, in: *Proceedings of ITC 14*, 1994, pp. 75–84.
- [14] S. Resnick, *Adventures in Stochastic Processes with Illustrations*, Birkhäuser, Boston, 1992.
- [15] M.F. Neuts, A versatile Markovian point process, *J. Applied Prob.* 16 (1979) 764–779.
- [16] F.A. Tobagi, T. Kwok, F.M. Chiussi, Architecture, performance, and implementation of the tandem banyan fast packet switch, *IEEE J. Select. Areas Commun.* 9 (8) (1991) 1173–1193.
- [17] A. Itoh, W. Takahashi, H. Nagano, M. Kurisaka, S. Iwasaki, Practical implementation and packaging technologies for a large-scale ATM switching system, *IEEE J. Select. Areas Commun.* 9 (8) (1991) 1280–1288.
- [18] F.A. Tobagi, Fast packet switch architectures for broadband integrated services digital networks, *Proc. IEEE* 78 (1) (1990) 133–167.
- [19] Y.S. Yeh, M.G. Hluchyj, A.S. Acampora, The knockout switch: a simple, modular architecture for high-performance packet switch in: *Proceedings of ISS'87*, 1987, pp. B10.2.1 to B10.2.8.
- [20] J.W. Roberts, J.T. Virtamo, The superposition of periodic cell arrival streams in an ATM multiplexer, *IEEE Trans. Commun.* 39 (2) (1991) 298–303.
- [21] D.M. Lucantoni, New results on the single server queue with a batch Markovian arrival process, *Commun. Statist. — Stochastic Models* 7 (1) (1991) 1–46.
- [22] A. Graham, *Kronecker Products and Matrix Calculus with Applications*, Ellis Horwood Series, Halsted Press, 1981.



Koohong Kang received B.Sc. and M.Sc. degrees in Electronics Engineering from Kyungpook National University and Chungnam National University, Korea, in 1985 and 1990, respectively. From August 1985 to February 1993, he was with the Electronics and Telecommunications Research Institute (ETRI), Korea, as a member of technical staff. He is currently a Ph.D. student in the Department of Computer Science and Engineering, Pohang University of Science and Technology (POSTECH), Korea. His research interests include performance analysis and traffic control in high-speed networks.



Cheeha Kim received his B.Sc. degree in Electronics Engineering from Seoul National University, Korea in 1974, and then M.Sc. and Ph.D. degrees from the University of Maryland, College Park, in 1984 and 1986, respectively, both in Computer Science. From 1974 to 1981, he was with the Agency for Defense Development (ADD), Korea, as a member of technical staff. From 1986 to 1989, Dr Kim was an Assistant Professor in the Department of Computer Science, State University of New York at Buffalo. Currently, Dr Kim is the Director of POSTECH Information Research Laboratories (PIRL), and an Associate Professor in the Department of Computer Science and Engineering, POSTECH, Korea. His research interests include computer networks, distributed systems, and performance evaluation.