Performance Evaluation 12 (1991) 181-189
North-Holland

Statistical multiplexing under non-i.i.d. packet arrival processes and different priority policies

Ioannis Stavrakakis
Department of Electrical Engineering and Computer Science, University of Vermont, Burlington, VT 05405, USA

Abstract

Stavrakakis, I., Statistical multiplexing under non-i.d.d. packet arrival processes and different priority policies, Performance Evaluation 12 (1991) 181-189.

In this paper some statistical multiplexing schemes under non-i.i.d. packet arrival processes are considered. Various packet multiplexing policies with priorities are proposed to introduce fairness in the service process, accommodate different packet delay requirements and avoid monopolization of the transmission media by some sources. The per input line packet arrival process is described as a Markov Modulated Generalized Bernoulli Process (MMGBP). The MMGBP can serve as a model for a wide class of complex packet arrival processes present in integrated services digital networks. Furthermore, when certain priority policies are in effect the original MMGBP can be transformed into another MMGBP where the priority policy is properly incorporated. As a result, auxiliary/equivalent FIFO multiplexing systems can be constructed with inputs described by a MMGBP, as well. Finally, the TDM application that is presented in this paper illustrates the appropriateness of the MMGBP in representing the effect of certain policies on the behavior of coupled (in some sense) queues. The above properties of the MMGBP facilitate the analysis of certain multiplexing systems under some dependent packet arrival processes.

Keywords: Communication networks, queueing systems, statistical multiplexing.

1. Introduction

Integrated Services Digital Networks (ISDNs) should not be seen as a simple evolution of Data Networks (DNs) which have been developed over the last two decades. The significant differences among the sources of information involved in ISDNs, regarding, for instance, the packet generation processes and the packet delivery requirements, create a more complex environment compared to that found in data networks.

Although the unit of information is a fixed size packet for all potential users of the system, to facilitate the integrating operation of an ISDN, the characteristics of the various packet processes of interest can be dramatically different from those present in a traditional data network. Poisson, Bernoulli, or general i.i.d. processes, widely incorporated in the analysis of data networks, are rather inappropriate for the description of the packet processes in an ISDN. For instance, packetized voice traffic can be modeled as blocks of packets arriving over consecutive time slots with geometrically distributed length (talkspurt) followed by periods of silence with geometrically distributed length. Other kinds of packetized information (such as long files, video traffic, etc.) may be described as blocks of packets whose length follows a general distribution. The output of a computer over a slot may contain more than one packet of information; fast transmission lines may also deliver more than one packet per slot. The packet traffic generated by a concentrator/transmitter and being delivered through a slotted line is constant (one packet per slot), whenever its buffer is non-empty and it is zero otherwise. Packet traffics generated by various sources in an ISDN or by network components in both an ISDN or a DN cannot be described with the memoryless models mentioned before.

In a discrete time slotted network, the packet traffics for the cases described above (among other ones) can be appropriately described by a Markov Modulated Generalized Bernoulli Process (MMGBP). That is, it is assumed that the source of information (i.e., network component or user) visits M states of an underlying Markov chain. Given the current state, the number of packets generated follows a general distribution. Clearly this packet process is a non-i.i.d. one. It is easy to establish that the cases of packet traffics described before may be described (or approximated) by a MMGBP. For instance, the packetized voice traffic is a MMGBP with two states, "talkspurt" and "silence". The probability that the voice source generates one packet when in state "talkspurt" is one; the probability that it generates zero packets when in state "silence" is one. The packet process of blocks of packets arriving over consecutive time slots may be described by a MMGBP [7] as well.

The other important issue in a packet network accommodating packets from sources with different characteristics is that of the allocation of the common facility among the sources. The allocation policy should take into consideration the time constraints imposed on certain packets and the possible monopolization of the common resource by certain sources over long periods; the latter could introduce unacceptable delays to short messages (e.g., consisted of single packets) of interactive communication or control information.

In this paper, we analyze a number of statistical multiplexing schemes under packet arrival processes described by a MMGBP and under various priority policies. The non-i.i.d. MMGBP may be appropriate for the description of complex packet processes, while the prioritization may introduce fairness and increased efficiency in the system.

A statistical multiplexer with N packet input processes, each of which is described by a MMGBP has been analyzed in [7], under the first-in-first-out (FIFO) service policy. The analysis of the system in [7] (the results of which are presented in the next section) is the ground on which the methodology for the analysis of the multiplexing schemes with priorities will be built.

Packets are assumed to arrive through slotted synchronous lines. That is, all packet arrivals are declared at common time instants which coincide with the end of the slots (slot boundaries). Discrete time queueing models for statistical multiplexing schemes under non-i.i.d. inputs and without priorities have been analyzed in the past [2,3,7–9]. Previous work on statistical multiplexing where packets with different priorities are involved, is heavily based on the assumption of a memoryless packet arrival process (e.g., Poisson) [1,5,6]. Notice that the proposed MMGBP includes simpler processes such as the Bernoulli or the generalized Bernoulli (more than one packet arrivals may occur over the same slot) processes and the first-order Markov process (arrival/no arrival), approximating packet arrivals in bursts or describing the packetized voice traffic. Even under these simple arrival processes and for the priority policies considered in this paper, the corresponding multiplexing schemes have not been analyzed before.

The rest of the paper is organized as follows. In the next section the statistical multiplexer presented in [7] is briefly described and the results from the analysis in [7] are presented. In Section 3, four different multiplexing schemes are considered and the methodology, based on the construction of systems equivalent to the one in [7], is presented. The mean buffer occupancy and the mean packet delay for all packet categories are derived for all cases considered. In Section 4, some numerical results on the mean packet delay

Ioannis Stavrakakis was born in Athens, Greece, in 1960. He received the Diploma in electrical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, in 1983 and the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, in 1988. In 1988 he joined the Department of Electrical Engineering and Computer Science, University of Vermont, Burlington, where he is presently an Assistant Professor. His research interests are in statistical communication theory, multiple-access algorithms, computer communication networks, queueing systems and system modeling and performance evaluation. Dr. Stavrakakis is a member of Tau Beta Pi, the Technical Chamber of Greece and the IEEE Communications Society.

are presented for the cases considered in Section 3. Finally, the conclusions of this work appear in the last section.

2. The FIFO statistical multiplexer

In this section we describe the statistical multiplexer analyzed in [7] and present the equations derived for the calculations of the mean buffer occupancy and the mean packet delay. This system will be modified to accommodate the priority policies in the next section. By establishing equivalent systems with the one presented briefly in this section, similar equations will be used for the derivation of the queueing results of interest in the next section.

A statistical multiplexer which is fed by N input lines is shown in Fig. 1. The input lines (which are mutually independent) are assumed to be slotted and packet arrivals and service completions are synchronized with the end of the slots. A slot is defined to be the fixed service (transmission) time required by a packet. At most one packet can be served in one slot. The first-in-first-out (FIFO) service discipline is adopted. Packets arriving at the same slot are served in a randomly chosen order. The buffer capacity is assumed to be infinite. The packet arrival process associated with line i is defined to be the discrete time process $\{a_j^i\}_{j\geq 0}$, $i = 1, 2, \ldots, N$, of the number of packets arriving at the end of the jth slot; $a_j^i = k$, $0 \leqslant k < \infty$, if k packets arrive at the end of the jth slot through input line i.

Let $\{z_j^i\}_{j\geq 0}$, be a finite state Markov chain imbedded at the end of the slots, which describes the state of the input line i. Let $S^i = \{x_0^i, x_1^i, \ldots, x_{M^i-1}^i\}$, $M^i < \infty$, be the state space of $\{z_j^i\}_{j\geq 0}$. It is assumed that the state of the underlying Markov chain determines (probabilistically) the packet arrival process of the corresponding line. That is, if $a^i(x^i): S^i \to Z_0$, is a probabilistic mapping from S^i into the non-negative finite integers, Z_0, then the probability that k packets arrive at the buffer at the end of the jth slot is given by $\phi(z_j^i, k) = \Pr\{a^i(z_j^i) = k\}$. Furthermore, it is assumed that there is at most one state, x_0^i such that $\phi(x_0^i, 0) > 0$ and that the rest of the states of the underlying Markov chain result in at least one (but a finite number of) packet arrivals, i.e., $\phi(x_k^i, 0) = 0$, for $1 \leqslant k \leqslant M^i - 1$. All packet arrivals are assumed to occur at the end of the slots. To avoid instability of the buffer queue it is assumed that there is always one state x_0^i, such as described above.

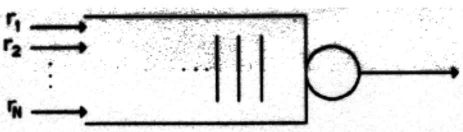

Fig. 1. The FIFO statistical multiplexer with N inputs.

The expected number of packets in the system is given by [7],

$$Q = \sum_{\bar{y}\in\bar{S}} W(\bar{y}) \tag{1}$$

where $\bar{S} = S^1 \times S^2 \times \ldots \times S^N$ and $W(\bar{y})$, $\bar{y} \in \bar{S}$, are the solutions of any $M^1 \times M^2 \times \ldots \times M^N - 1$ linear equations given by

$$W(\bar{y}) = \sum_{\bar{x}\in\bar{S}} W(\bar{x}) p(\bar{x}, \bar{y}) + \sum_{\bar{x}\in\bar{S}} (\mu_{\bar{x}} - 1) p(\bar{x}, \bar{y}) \pi(\bar{x}) + \sum_{\bar{x}\in\bar{S}} q_0(\bar{x}) p(\bar{x}, \bar{y}), \quad \bar{y} \in \bar{S} \tag{2a}$$

and the linearly independent equation

$$\sum_{\bar{x}\in\bar{S}} \left[2(\mu_{\bar{x}} - 1) W(\bar{x}) + 2(\mu_{\bar{x}} - 1) q_0(\bar{x}) + (2 + \sigma_{\bar{x}} - 3\mu_{\bar{x}}) \pi(\bar{x})\right] = 0 \tag{2b}$$

where

$$\pi(\bar{x}) = \prod_{i=1}^{N} \pi^i(x^i), \qquad p(\bar{x}, \bar{y}) = \prod_{i=1}^{N} p^i(x^i, y^i),$$

$$q_0(\bar{x}) = (1 - \lambda) p(\bar{x}_0, \bar{x}),$$

$$\mu_{\bar{x}} = \sum_{\nu=1}^{R} \nu g_{\bar{x}}(\nu), \qquad \sigma_{\bar{x}} = \sum_{\nu=1}^{R} \nu^2 g_{\bar{x}}(\nu),$$

$$g_{\bar{x}}(\nu) = Pr\left\{\sum_{i=1}^{N} a^i(x^i) = \nu\right\}$$

and

$$\lambda = \sum_{\bar{x}\in\bar{S}} \mu_{\bar{x}} \pi(\bar{x}) < 1$$

is the total input traffic which is less than 1 for stability. R is the maximum number of packets which may arrive at the same slot from all N lines; $\pi^i(x^i)$ and $p^i(x^i, y^i)$ are the steady state and the transition probabilities of the Markov chain associated with the ith input line. The mean packet delay is given by using Little's formula, i.e.,

$$D = Q/\lambda \tag{3}$$

3. Statistical multiplexing with priorities

In this section we consider various multiplexing schemes under different priority policies. The per slot and line packet arrival processes are described by the MMGBP introduced in Section 2.

3.1. Case 1

Consider the statistical multiplexer shown in Fig. 2; the input lines, r_1 and r_2, are assumed to carry synchronous packet traffic. The packet arrival processes $\{a_j^1\}_{j\geq 0}$ and $\{a_j^2\}_{j\geq 0}$ are assumed to be two MMGBPs. In particular, $\{a_j^1\}_{j\geq 0}$ is assumed to be a MMGBP with two underlying states x_0^1 and x_1^1 and packet generation probabilities $\phi^1(x_0^1, 0) = 1$ and $\phi^1(x_1^1, 1) = 1$. That is, one packet is generated when the line (or the source connected to the line) is in state x_1^1 and no packet is generated when in state x_0^1. This model may describe the packet traffic generated by a voice source or, in general, blocks of packets of geometrically distributed length, arriving over consecutive slots. The second packet process $\{a_j^2\}_{j\geq 0}$ is assumed to be given by the general MMGBP described in the previous section.

In the statistical multiplexing scheme considered here it is assumed that line r_1 carries high priority traffic which has priority over that carried by line r_2. That is, it is assumed that the server (which makes decisions at the slot boundaries) moves to line r_2 only if the buffer associated with line r_1 is empty; it returns to line r_1 as soon as the corresponding buffer associated with line r_1 becomes non-empty. Since at most one packet arrives through line r_1, the service policy implies that a single packet buffer is required for line r_1. If the cut-through connection is possible, no buffer is necessary for line r_1. An infinite capacity buffer is assigned to line r_2.

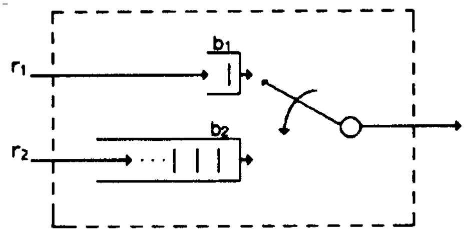

Fig. 2. The statistical multiplexer of Case 1.

Clearly, there are two categories of packets, say C_1 and C_2, with different priorities (a smaller subscript indicates higher priority). Packets in C_1 are served (transmitted) right away. Thus, the mean delay of packets in C_1, D_1, is equal to 1 (the service time). Service of packets in C_2 is interrupted whenever a packet arrives through line r_1; let D_2 be the mean delay of packets in C_2.

To compute D_2 we consider a FIFO system (shown in Fig. 1) which is equivalent to the one considered here. An *equivalent FIFO system* is defined as a FIFO system whose packet arrival processes are identical to those of the system under consideration; let D_{12} denote the mean packet delay induced by the equivalent FIFO system. Since the queueing system is work conserving and nonpreemptive, the conservation law [1,4] implies that D_{12} satisfies the following equation.

$$D_{12} = \frac{\lambda_1 D_1 + \lambda_2 D_2}{\lambda_1 + \lambda_2} \tag{4}$$

where λ_1 and λ_2 are the per slot packet arrival rates through lines r_1 and r_2, respectively. D_{12} can be computed from eqns. (1)–(3). Then D_2, the mean delay of packets in C_2, can be computed from (4) by setting $D_1 = 1$.

A practical application of the simple priority scheme described here is related to the mixing of voice and data packets; r_1 may carry packetized voice ($\lambda_1 < 0.5$) and r_2 may carry blocks of packets of time unconstrained information. The multiplexing scheme provides (in essence) a circuit to the voice traffic which is utilized by data packets when idle. The mean data packet delay, in this case, is given by D_2.

Another application of the above priority scheme, which is of both theoretical and practical significance, is related to the analysis of a Time Divison Multiplexer (TDM) under station traffic described by a MMGBP. This application illustrates the capability of the MMGBP in describing imaginary packet processes which represent the operation of a system. At the same time mean delay results for the particular multiplexer are easily derived.

Consider a time division multiplexing system with N buffered stations. Each station is assigned one slot per frame; the frame is supposed to be consisted of N slots. The per station packet arrival processes are assumed to be modeled as MMGBPs. Although the queues of the stations do not interfere directly with each other, the service policy introduces a (deterministic) coupling among the queues, in the sense that the presence of the other $N-1$ queues (users) results in a service policy which removes one packet from the queue under study (if nonempty) every N slots. It is the number of queues (users) in the system and not their status that, in conjunction with the service policy, introduces the coupling which makes the analysis difficult.

To study the queueing system associated with, for example, station 1, a second packet arrival process (input line) to its buffer is considered for the representation of the coupling. This process, denoted by $\{\tilde{a}_j^1\}_{j\geq 0}$, can be modeled as a MMGBP. The corresponding underlying Markov chain has N states, denoted by $1, 2, \ldots, N$, and transition probabilities given by

$$p(k, j) = \begin{cases} 1 & j = k+1, 1 \leq k < N \\ 1 & j = 1, k = N \\ 0 & \text{otherwise} \end{cases}$$

The corresponding probabilistic mapping is given by: $\tilde{a}^1(1) = 0$ and $\tilde{a}^1(\mathrm{k}) = 1$ for $1 < k \leq N$, with probability one.

From the above construction of the arrival process $\{\tilde{a}_j^1\}_{j\geq 0}$ it turns out that one packet arrives through the second line in every slot except from the first of a sequency of N consecutive slots. By assuming priority for these packets the decoupling of the queue under study is achieved. Whenever the server of the TDM system serves the other stations, the server of the decoupled queueing system serves the priority packets arriving from the process $\{\tilde{a}_j^1\}_{j\geq 0}$. Thus, the time division multiplexing policy of the original system is represented by the second packet arrival process $\{\tilde{a}_j^1\}_{j\geq 0}$ to the queue under study. Clearly, the queueing system in the TDM station is identical to that of Fig. 2 where line r_1 carries the traffic from the process $\{\tilde{a}_j^1\}_{j\geq 0}$ and line r_2 carries the traffic to the station under study. The desired mean packet delay of the particular station corresponds to D_2 and can be computed as described above.

3.2. Case 2

Consider the statistical multiplexer shown in Fig. 3. Both synchronous traffics $\{\alpha_j^1\}_{j\geq 0}$ and $\{\alpha_j^2\}_{j\geq 0}$ are assumed to be modeled as MMGBPs. Case 2 is identical to Case 1 with the only difference being that more than one packets per slot may arrive through line r_1, as well. As a result, queueing problems appear in both lines. Line r_1 carries high priority traffic (or the source connected to r_1 has priority over the one connected to line r_2) which has priority over that carried by line r_2. To compute D_1 and D_2, in this case, we proceed as follows.

Calculation of D_1

Consider a FIFO statistical multiplexer with one input line which is identical to r_1. By using eqns. (1)–(3), we compute the mean packet delay induced by this FIFO multiplexer. Clearly, this mean packet delay is equal to D_1. The priority of r_1 over r_2 results in a buffer behavior of line r_1 which is not affected by the packet arrival process in r_2. Thus, the behavior of the buffer connected to r_1 is identical to that of the FIFO multiplexer described above.

Calculation of D_2

To compute the mean delay of packets in C_2 we use the equivalent FIFO statistical multiplexer. The mean packet delay, D_{12}, is obtained from eqns. (1)–(3). Then D_2 is obtained from (4).

3.3. Case 3

Consider the statistical multiplexer shown in Fig. 4. The packet arrival process $\{a_j^1\}_{j\geq 0}$ is assumed to be a MMGBP, as described in Section 2. To avoid monopolization of the facility by long messages (consisted of many packets) which arrive over a single slot, the following service policy is

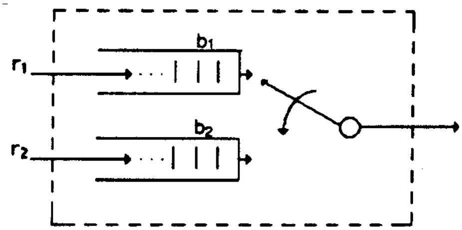

Fig. 3. The statistical multiplexer of Case 2.

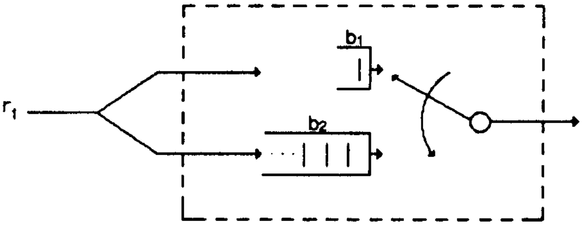

Fig. 4. The statistical multiplexer of Case 3.

introduced. The first packet of those arriving during a single slot enters a single packet buffer b_1 and it is transmitted in the next slot. The rest of the packets enter an infinite capacity buffer b_2. The server moves to buffer b_2 only if buffer b_1 is empty. This service discipline gives priority to single packets (over a slot); packets other than the first of a slot are served under a FIFO policy interrupted by new arrivals. This service policy introduces some fairness in the service policy and favors single packets.

Clearly, the mean delay of single packets (or of the first packet of a multipacket of a slot) is equal to 1 slot, i.e. $D_1 = 1$. The mean delay of packets which enter b_2 is given by (4), where λ_1 is equal to $\pi(x \neq x_0)$ (the probability that the line is in any of the packet generating states), $\lambda_2 = \lambda_{\text{total}} - \lambda_1$ and D_{12} is the mean packet delay of the equivalent FIFO multiplexer of Fig. 1 computed from eqns. (1)–(3).

3.4. Case 4

Consider the statistical multiplexer shown in Fig. 5. The per input line packet arrival process and the service policy are as in Case 3. The first packet per slot arriving in each of the input lines is given priority by being sent to the infinite capacity buffer b_1; the rest of the packets arriving over the same slot are sent to the infinite buffer b_2. The FIFO service policy is assumed for the packets of the same buffer. Packets in b_1 have priority over those in buffer b_2. That is, service of the packets in b_2 can start only if buffer b_1 is empty. This service policy avoids monopolization of the facility by either long messages (independently of the generating source) or certain sources (which by nature generate long messages). To compute D_1 and D_2 we proceed as follows.

Calculation of D_l

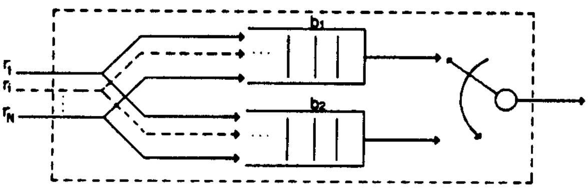

Fig. 5. The statistical multiplexer of Case 4.

Consider the statistical multiplexer shown in Fig. 5. The per input line packet arrival process and the service policy are as in Case 3. The first packet per slot arriving in each of the input lines is given priority by being sent to the infinite capacity buffer b_1; the rest of packets arriving over the same slot are sent to the infinite buffer b_2. The FIFO service policy is assumed for the packets of the same buffer. Packets in b_1 have priority over those in buffer b_2. That is, service of the packets in b_2 can start only if buffer b_1 is empty. This service policy avoids monopolization of the facility by either long messages (independently of the generating source) or certain sources (which by nature generate long messages). To compute D_1 and D_2 we proceed as follows.

Calculation of D_l

Consider a FIFO statistical multiplexer (Fig. 1) whose packet arrival process is given by MMGBPs. The underlying Markov chains of these MMGBPs are identical to those associated with the input lines $r_1, \ldots, r_N$. The probabilistic mapping

$$a(\bar{x}) = \sum_{i=1}^{N} a^i(x^i), \quad \bar{x} \in \bar{S}$$

is modified to describe the packet arrival process to b_1. That is,

$$a_1(\bar{x}) = \sum_{i=1}^{N} 1_{\{x^i \neq x_0^i\}}, \quad \bar{x} \in \bar{S} \tag{5}$$

where x_0^i is the state of line i which generates no packets. Based on (5), the packet generating probabilities $\phi^i(x^i, k)$ are modified to the following

$$\phi^i(x_0^i, 0) = 1 \quad \text{and } \phi^i(x^i, 1) = 1 \text{ for } x^i \neq x_0^i \tag{6}$$

The mean delay of the packets in D_1 is now computed by applying eqns. (1)–(3) on the FIFO system with packet arrival processes as determined by (6). The total packet arrival rate λ (used in (3)) is given by

$$\lambda_1 = \sum_{i=1}^{N} \pi\left(x^i \neq x_0^i\right). \tag{7}$$

Calculation of D_2

The mean delay of packets in buffer b_2 is computed from (4), where λ_1 is given by (7), $\lambda_2 = \lambda_{\text{total}} - \lambda_1$, and D_{12} is the mean packet of the equivalent FIFO multiplexer of Fig. 1, computed from eqns. (1)–(3).

Table 1
Mean packet delay results for Case 1.a

λ	γ	D_1	D_{12}	D_2
0.90	0.5	1.000	13.897	26.794
0.90	0.3	1.000	9.325	17.651
0.90	0.0	1.000	5.897	10.794
0.70	0.5	1.000	4.799	8.598
0.70	0.3	1.000	3.466	5.981
0.70	0.0	1.000	2.466	3.931

4. Numerical results

In this section some numerical results are derived for each of the four priority policies described in the previous section. In the examples considered below it is assumed that the underlying Markov chain associated with any of the input lines has two states, that is $S^i = \{0, 1\}$ for the ith line. State 0 is the no-packet generating state (i.e., $a^i(0) = 0$); state 1 generates at least one packet, up to a maximum of K^i, with probabilities $\phi^i(1, j)$, $1 \leqslant j \leqslant K^i$.

As the delay results illustrate, an input traffic process which generates packets clustered around consecutive slots and followed by a period of inactivity, causes significant queueing problems and the induced packet delay is greater that the one induced under better randomized packet arrivals of the same intensity. Since state 1 generates packets and state 0 does not, it makes sense to use the quantity γ^i, where,

$$\gamma^i = p^i(1, 1) - p^i(0, 1) \tag{8}$$

as a measure of the clusterness of the packet arrival traffic; $p^i(k, j)$ is the probability that the Markov chain associated with line i moves from state k to state j. The value of $\gamma^i = 0$ corresponds to a per slot independent packet generation process (generalized Bernoulli process). The clusterness coefficient γ^i and the packet arrival rate λ^i are two important quantities which dramatically affect the delay induced by the multiplexing system. For this reason, each traffic will be characterized by the pair (λ^i, γ^i) and the distribution $\phi^i(1, j)$, $1 \leqslant j \leqslant k^i$. The rest of the parameters of the MMGBPs associated with each input line are computed from the following equations:

$$\pi^i(1) = \frac{\lambda^i}{\sum_{j=1}^{K^i} j\phi^i(1, j)}$$

$$\pi^i(0) = 1 - \pi^i(1) \tag{9a}$$

$$p^i(0, 1) = (1 - \gamma^i)\pi^i(1),$$
$$p^i(1, 1) = \gamma^i + p^i(0, 1) \tag{9b}$$

$$p^i(1, 0) = 1 - p^i(1, 1),$$
$$p^i(0, 0) = 1 - p^i(0, 1) \tag{9c}$$

4.1. *Case 1*

Consider the multiplexing system of Case 1 with distributions $\phi^1(1, 1) = 1$, $\phi^2(1, 1) = 0.5$, $\phi^2(1, 2) = 0.3$, $\phi^2(1, 3) = 0.2$ and parameters $\lambda^1 = \lambda^2 = \lambda/2$ and $\gamma^1 = \gamma^2 = \gamma$ (Case 1.a). The mean packet delay results D_1, D_2 and D_{12} are given in Table 1, for various values of λ and γ. It can be easily observed that for a given total input rate λ,

Table 2
Mean packet delay results for Cases 1.b and 1.c

λ^2	γ^2	Case 1.b		Case 1.c	
		D_{12}	D_2	D_{12}	D_2
0.55	0.5	41.207	66.794	33.694	54.500
0.55	0.3	37.541	60.794	32.472	52.500
0.55	0.0	34.790	56.294	31.139	51.000
0.35	0.5	11.966	22.931	9.917	18.833
0.35	0.3	10.966	20.931	9.583	18.167
0.35	0.0	10.216	19.431	9.333	17.667

Table 3
Mean packet delay results for Case 1.d

λ	$\gamma=0.0$	$\gamma=0.3$	$\gamma=0.5$
0.04	8.500	14.928	23.500
0.06	12.250	21.893	34.750
0.08	23.500	42.785	68.500

the smallest induced delay is achieved for $\gamma=0$ (independent per slot packet generation process). This is due to the fact that $\gamma=0$ results in the best randomization of the packet arrivals for given λ and $\phi^1(1, j)$, $0 \leqslant j \leqslant K^i$.

When $\lambda^1=0.35$ and $\gamma^1=0.93$, line 1 may describe packetized voice traffic with geometrically distributed talkspurt periods (with mean ~22 packets) and geometrically distributed silence periods (with mean ~40 packets) [3]. The distributions of $\phi^1(1, 1)$ and $\phi^2(1, j)$, $1 \leqslant j \leqslant 3$, are the same as before. The mean delay results are shown in Table 2 for various values of λ^2 and γ^2 (Case 1.b). Notice that although the total traffics considered are equal to those in Table 1, the induced mean packet delay D_2 is much larger, due to the larger value of the clusterness coefficient γ^1.

For $\lambda^1=0.35$, $\gamma^1=0.93$ and $\phi^2(1,1)=1$, the induced mean packet delay D_2 is smaller than that of Case 1.b, for the same values of λ^1, γ^1, λ^2 and γ^2 (Case 1.c). This is due to the reduced clusterness resulting from the fact that only single packets arrive through line 2, as well (as opposed to possibly multiple packets arriving under the previous case). These results are shown in Table 2 (Case 1.c).

Finally, consider a TDM system with $N=10$ stations and packet arrival process to the station under study given by a 2-state MMGBP with parameters $\phi(0,0)=1$, $\phi(1,1)=0.5$, $\phi(1,2)=0.3$, $\phi(1,3)=0.2$ (Case 1.d). By following the analysis approach described in Section 3, the mean packet delay is calculated; the results are shown on Table 3, for various values of λ and γ. These results indicate that the presence of memory in the packet arrival process (as captured by γ) has a tremendous effect on the resulting induced packet delay. For instance, if a packet arrival process with parameters $\lambda=0.06$ and $\gamma=0.3$ is approximated by an independent process ($\gamma=0.0$) with the same arrival rate, the obtained delay result is equal to 12.250 slots as opposed to the accurate 21.893 slots.

4.2. Case 2

Consider the multiplexing system of Case 2 with probability distribution $\phi^1(1, 1)=0.6$, $\phi^1(1, 2)=0.4$, $\phi^2(1, 2)=0.3$, $\phi^2(1, 4)=0.5$, $\phi^2(1, 6)=0.2$ and the parameters $\lambda^1=\lambda^2=\lambda/2$ and $\gamma^1=\gamma^2=\gamma$. The mean packet delay results D_1, D_2 and D_{12} are shown in Table 4 for various values of λ and γ. Notice that $D_1>1$ since more than one packets may arrive over the same slot through line 1.

4.3. Case 3

Consider the multiplexing system of Case 3 with probability distribution $\phi^1(1, 1)=0.4$, $\phi^1(1, 2)=0.3$, $\phi^1(1, 3)=0.2$, $\phi^1(1, 4)=0.1$. The mean packet delay results D_2 and D_{12} are shown in Table 4 for various values of $\lambda^1=\lambda$ and $\gamma^1=\gamma$.

4.4. Case 4

Consider the multiplexing system of Case 4 with $N=3$ input lines, probability distributions as in Case 3 and parameters $\lambda^1=\lambda^2=\lambda^3=\lambda/3$ and $\gamma^1=\gamma^2=\gamma^3=\gamma$. The mean packet delay results

Table 4
Mean packet delay results for Cases 2, 3 and 4

λ	γ	Case 2			Case 3		Case 4		
		D_1	D_{12}	D_2	D_{12}	D_2	D_1	D_{12}	D_2
0.9	0.5	2.247	33.468	64.689	18.500	36.000	1.818	27.500	53.181
0.9	0.3	1.831	21.754	41.676	12.786	24.571	1.506	18.357	35.208
0.9	0.0	1.519	12.968	24.416	8.500	16.000	1.273	11.500	21.727
0.7	0.5	2.055	11.323	20.590	6.833	12.557	1.538	9.167	16.795
0.7	0.3	1.703	7.608	13.513	4.928	8.857	1.333	6.373	11.413
0.7	0.0	1.439	4.823	8.206	3.500	6.000	1.179	4.278	7.376

are shown in Table 4 for various values of λ and γ.

References

[1] G. Barberis, A useful tool in the theory of priority queueing, *IEEE Trans. Comm.* **28** (9) (1980) 1757–1762.

[2] H. Bruneel, Queueing behavior of statistical multiplexers with correlated inputs, *IEEE Trans. Comm.* **36** (12) (1988) 1339–1341.

[3] H. Heffes and D. Lucantoni, A Markov modulated characterization of packetized voice and data traffic and related statistical multiplexer analysis, *IEEE J. Selected Areas Comm.* **4** (6) (1986) 856–868.

[4] D. Heyman and M. Sobel, *Stochastic Models in Operations Research, Vol. 1* (McGraw-Hill, 1982).

[5] M. Hluchyj, C. Tsao and R. Boorstyn, Performance analysis of a preemptive priority queue with applications to packet communication systems, *Bell System Tech. J.* **62** (10) (1983) 3225–3245.

[6] C.-Y. Lo, Performance analysis and application of a two priority packet queue, *AT&T Tech. J.* **66** (3) (1987) 82–89.

[7] I. Stavrakakis, Analysis of statistical multiplexer under a general input traffic model, IEEE INFOCOM '90, June 5–7, 1990, San Francisco, pp. 1220–1225.

[8] D. Towsley, The analysis of a statistical multiplexer with nonindependent arrivals and errors, *IEEE Trans. Comm.* **28** (1) (1980) 65–72.

[9] A. Viterbi, Approximate analysis of time synchronous packet networks *IEEE J. Selected Areas Comm.* **4** (6) (1986) 879–890.