IEEE

A Time Division Multiplexer With Dependent Packet Arrival Processes

Ioannis Stavrakakis
EE/CS Department
University of Vermont
Burlington, Vermont 05405

Abstract

In this paper a Time Division Multiplexer (TDM) is analyzed under a general model for the per station packet arrival process. This model can be suitable for the description of the packet generation process of a variety of packet sources. The expected packet delay of each station is exactly derived. The obtained results show that the mean packet delay may vary significantly for different stations, as long as the statistical characteristics of the packet traffics are different, even if the intensity of the packet traffics and the assigned capacities are the same. The latter implies that oversimplifications in the description of the packet traffics may result in very inaccurate results.

I. Introduction

We consider a multiple access communication channel which is intended to serve a population of N (N≥2) stations, each wishing to transmit its messages across the channel. A satellite community of earth stations, a terrestrial radio channel providing the communication medium to a number of data terminals, a local area communications network, or a multiplexed link in a computer network may serve as examples, [1].

The sharing of the channel by the network users (stations) is supervised and controlled by the underlying access control discipline. A multitude of access-control schemes have been devised and studied. For instance, one can distinguish among procedures involving fixed assignment schemes, reservation schemes, polling schemes, random access schemes, or procedures which integrate several of the above mentioned schemes.

In this paper, a TDMA (time division multiple access) fixed channel assignment policy will be considered. Under a TDMA scheme each station is assigned, on a fixed basis, channel transmission time. Time is divided into successive periods of constant duration called (time) frames. Each frame is subdivided into M (M≥N) successive slots. More than one slots over the same frame may be assigned to a station.

The difference between the TDMA systems analyzed in the past, [2]-[4], and the system considered here is in the adopted models for the message arrival traffic. The vast majority of the TDMA systems analyzed in the past assumes a Poisson (or a batch Poisson) per station message arrival process. The lack of memory in this process facilitates the analysis of the formulated queueing systems. The Poisson model is inaccurate for the description of packet processes as generated and/or processed by elements of today's complex packet communication networking structures, due to the dependencies introduced by those elements. For instance, a message (consisted of K packets) transmitted over the network lines will arrive at a TDMA station over K consecutive slots (packet transmission time over fixed speed transmission lines) rather than as a batch process. If the message is affected by network routing decisions and multiplexing procedures, then it is most probable that the K particular packets will arrive at the TDMA station neither over the same slot (a batch) nor over K consecutive slots.

The paper is organized as follows. In the next section, some packet processes appearing in packet communication networks are presented and a common model for their unified description is developed. In section III, the TDMA system to be studied is described under per station packet arrival processes given by the general model for the dependent packet processes presented in secton II. This system is analyzed by invoking the results from [5]. Finally, some numerical results are presented in the last section.

II. A model for some dependent packet processes.

The description of the dependent packet process generated by network elements is essential to the analysis of the formulated queueing systems. In this work, dependent packet processes will be described by the following Markov Modulated Generalized Bernoulli (MMGB) model.

Definition

Assume that a network element which generates packet traffic satisfies the following:

(a) There exists an ergodic Markov chain $\{z_j\}_{j\geq 0}$ associated with the description of the state of the element; let $S=\{x_1,x_2,\ldots,x_M\}$, $M<\infty$, be the state space of $\{z_j\}_{j\geq 0}$ and $p(x_k,x_l)$, $\pi(x_k)$, $x_k,x_l \in S$, be the corresponding state transition and steady state probabilities.

(b) There exists a stationary probabilistic mapping $a(z_j)$: $S \to Z_0$ (where Z_0 is the set of nonnegative finite integers), which describes the number of packets departing at the end of the j^{th} time interval (slot). Let $a(z_j)=\rho$, $0\leq\rho\leq\infty$, $z_j \in S$, with probability $\phi_\rho(z_j)$.

Then, the packet process generated by the network element is given by

$$\{a_j\}_{j\geq 0}=\{a_j(z_j)\}_{j\geq 0} \qquad (1)$$

i.e, it is described as a Markov modulated generalized Bernoulli process. A 2-state Markov modulated Poisson process has been adopted in [6] for the approximate characterization of the superposition of voice and data traffic.

Notice that the process $\{a_j\}_{j\geq 0}$, as given by (1), describes exactly the output process of a network element, provided that the conditions in the Definition are satisfied. Some examples of network elements generating dependent packet processes are the following:

Example 1: Bursty traffic network links.

Consider a link which carries traffic modulated by various other components of a large network and by routing decisions. The network component in this case is the link and its input and output processes are identical. In [7] it has been found that network packet traffic is bursty. As a result, a first order Markov model has been adopted for the description of this packet process. If p(0,1) and p(1,1) are the conditional probabilities of a packet arrival (departure) given that 0 or 1 arrivals (departures) occurred in the previous slot, respectively, then the burstiness coefficient is defined, [7], as

$$\gamma = p(1,1) - p(0,1)$$

This traffic model can be easily described in terms of the proposed model on the packet process generated by the link (network component). Let z_j be the number of packets in the link at the beginning of the j^{th} slot. Clearly, $\{z_j\}_{j\geq 0}$ is a Markov chain with S={0,1}. The transition probabilities are identical to those of the first order Markov model that describes the bursty traffic. The mapping $a(z_j)$ is, in this case, deterministic and it is given by

$$\phi_1(0)=0 \ , \ \phi_1(1)=1 \ , \ \phi_0(0)=1 \ , \ \phi_0(1)=0$$

The packet (output) process of the link just described, can be the model for an on/off switch with Bernoulli packet arrivals. In this case, when the switch is on the output process is the same Bernoulli process and it is zero otherwise. Markov models can usually be incorporated in the description of the on-off activity of a switch. A switch in the off position could correspond to a failure or to a situation in which it serves other links.

The parameters of the packet output process of a network component which generates bursty traffic are determined from the packet rate $\pi(1)$ and the burstiness coefficient γ. Given these quantities, the rest of the parameters of the Markov model are calculated from the equations

$$\pi(0) = 1-\pi(1) \ , \ p(0,0) = 1-p(0,1) \ , \ p(1,0) = 1-p(1,1)$$

$$p(0,1) = \pi(1)(1-\gamma) \ , \ p(1,1) = \gamma + p(0,1).$$

Example 2: The single message node

Consider now a network node which is capable of storing and forwarding a single message at a time. It is assumed that the input process to this component is Bernoulli with intensity μ messages per slot. Each message is assumed to consist of a variable number of packets; let $\sigma(i)$ = Pr{message consists of i packets}, $1\leq i\leq K$. The single message buffering assumption implies that messages which find the component non-empty are either discarded or served by a (buffered) low priority link. In the second scenario, the link served by the node under consideration is reserved for new messages. These messages are given a chance to be transmitted right away (if the line is not busy), before they enter a (probably) first in first out queue formulated in the input of another link. Without loss of generality, it is assumed that a new message is also accepted if there is only one packet (the last of the previous message) in the node. It is assumed that arrivals occur at the beginning of a slot. As a result, a new message may start being served right after the end of the previous message transmission. The packet output process of this component is definitely a non-Bernoulli process. It can be easily described in terms of the processes $\{z_j\}_{j\geq 0}$ and $\{a_j\}_{j\geq 0}$ defined in the previous section. If z_j is the number of packets in the node at the end of the j^{th} slot, then it can be easily shown that $\{z_j\}_{j\geq 0}$ is a Markov chain with state space $S=\{0,1,2,\cdots,K\}$. The transition probabilities are given by

$$p(0,i) = p(1,i) = \mu\sigma(i) \ , \ 1\leq i\leq K$$

$$p(0,0) = p(1,0) = 1-\mu$$

$$p(k,k-1) = 1 \ , \ 2\leq k\leq K$$

$$p(k,i) = 0 \ , \ 2\leq k\leq K \ , \ 1\leq i\leq K \ , \ i=k-1$$

Given μ and $\sigma(i)$, $1\leq i\leq K$, the steady state probabilities $\pi(i)$, $0\leq i\leq K$ can be easily computed. The mapping given by (1) is deterministic in this case and has the following parameters

$$\phi_1(k) = 1 \ , \ 1\leq k\leq K \ , \ \phi_1(0) = 0$$

$$\phi_0(k) = 1-\phi_1(k) \ , \ 0\leq k\leq K,$$

A Bernoulli approximate model on the output process of the node would have intensity

$$\lambda = \sum_{i=1}^{K}\pi(i) = 1-\pi(0)$$

A better approximate model on the true packet output process could be a first order Markov model. If 1 and 0 denote one or zero packet outputs, respectively, then the parameters of this Markov model are given by

$$\pi_m(0) = \pi(0) \ , \ \pi_m(1) = 1-\pi_m(0)$$

$$p_m(0,0) = 1-p_m(0,1) \ , \ p_m(1,0) = 1-p_m(1,1)$$

$$p_m(0,1) = \mu \ , \ p_m(1,1) = 1-p_m(0,1)\frac{\pi_m(0)}{\pi_m(1)}$$

and the corresponding burstiness coefficient γ is given by

$$\gamma = p_m(1,1) - p_m(0,1)$$

Example 3: A node with arbitrarily large buffer.

In this case it is assumed that all messages which would fit into the buffer of size $M<\infty$ are received; no message is partially received. Let g(k), $0\leq k\leq K$, be the probability that a message with k packets arrives over a slot; k=0 corresponds to no message arrival.

Similarly to the previous example, the output process of the finite buffer node can be easily described in terms of the processes $\{z_j\}_{j\geq 0}$ and $\{a_j\}_{j\geq 0}$. If z_j is the number of packets in the node at the end of the j^{th} slot, then $\{z_j\}_{j\geq 0}$ is a Markov chain with state space $S = \{0,1,2,\ldots,M\}$. The transition probabilities are given by (assume g(k)=0 for k>K)

$$p(0,j) = g(j) \ , \ 0\leq j\leq M$$

$$p(k,j) = g(j-k+1) \ , \ 1\leq k\leq M, \ k-1\leq j\leq M$$

and the probabilistic mapping is determined by

$$\phi_1(k) = 1 \ , \ 1\leq k\leq M \ , \ \phi_1(0) = 0$$

$$\phi_0(k) = 1-\phi_1(k) \ , \ 0\leq k\leq M$$

The Bernoulli and the Markov approximations on the resulting packet output traffic can be determined as in the previous example.

III. Time division multiplexing

Consider the time division multiplexing system shown in Fig. 1. Each of the N buffered users 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 MMGB processes. Although the queues of the stations do not interfere directly with each other, the service policy introduces a (deterministic) coupling to each of 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.

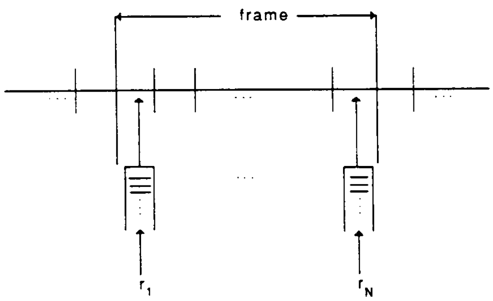

Figure 1
A TDMA communication system with N stations.

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

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

The corresponding probabilistic mapping is given by: $a(1)=0$ and $a(k)=1$ for $1<k\leq M$, with probability one.

From the above construction of the arrival process $\{\bar{a}_j^1\}_{j\geq 0}$ it turns out that one packet arrives through the second line in every slot except from the first of a sequence of N consecutive slots. By assuming preemptive priority for these packets the decoupling of the queue under study is achieved. Whenever the server of the TDMA system serves the other users the server of the decoupled queueing system serves the preemptive priority packets arriving through line $\{\bar{a}_j^1\}_{j\geq 0}$. Thus, the time division multiplexing policy of the original system is represented by the second packet arrival process $\{\bar{a}_j^1\}_{j\geq 0}$ to the queue under study.

The mean packet delay, D_{12}, induced by the decoupled queue with arrival processes $\{a_i^1\}_{j\geq 0}$ and $\{\bar{a}_j^1\}_{j\geq 0}$ can be computed by considering the equivalent FCFS (First-Come First-Served) system (Fig. 2), which has been analyzed in [5]. The description of this multiplexer and the analysis results are briefly presented in a companion paper. Then, the mean packet delay, D_2, in the buffer of user 1 can be computed from the expression:

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

D_1 is the mean delay of the packets generated by $\{\bar{a}_j^1\}_{j\geq 0}$ and it is equal to 1; λ_1 is the packet arrival rate of the process $\{\bar{a}_j^1\}_{j\geq 0}$ and it is equal to (N-1)/N; λ_2 is the packet arrival rate of the process $\{a_i^1\}_{j\geq 0}$.

IV. Numerical results

In this section some numerical results for the mean packet delay induced by the TDMA system described in the previous section, are derived. The per station packet arrival processes are assumed to be described by a MMGB model which is based on a two state underlying Markov chain with state space $S^i=\{0,1\}$. 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 $\phi_1^i(j)$, $1\leq j\leq K^i$ (superscript i refers to the ith arrival process).

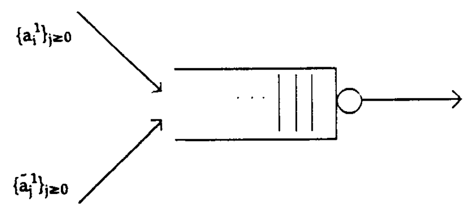

Figure 2
The decoupled buffer of the station under study.

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. Since state 1 generates packets and state 0 does not, it makes sense to use the quantity γ^i , where,

$$\gamma^i=p^i(1,1)-p^i(0,1)$$

as a measure of the clusterness of the packet arrival traffic; γ^i could also be seen as a measure of the intensity of the coupling in time; $p^i(k,j)$ 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 queueing system. For this reason, each traffic will be characterized by the pair (λ^i,γ^i) and the distribution $\phi_1^i(j)$, $1\leq j\leq k^i$. The rest of the parameters of the MMGB processes associated with each input line are computed from the following equations:

$$\pi^i(1) = \frac{\lambda^i}{\sum_{j=1}^{K^i} j\phi_1^i(j)} \quad , \quad \pi^i(0) = 1 - \pi^i(1)$$

$$p^i(0,1) = (1-\gamma^i)\pi^i(1) \quad , \quad p^i(1,1) = \gamma + p^i(0,1)$$

$$p^i(0,0) = 1-p^i(0,1) \quad , \quad p^i(1,0) = 1-p^i(1,1)$$

When N=10 and the packet arrival process to the station under study is given by a 2-state MMGB process (as described before) with parameters $\phi_1^1(1)=1$, $\phi_1^2(1)=.5$, $\phi_1^2(2)=.3$, $\phi_1^2(3)=.2$, the mean packet delay results are given in Table I, 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 $\lambda=.06$ and $\gamma=.3$ is approximated by an independent process ($\gamma=.0$) with the same arrival rate, the obtained delay result is equal to 12.250 slots as opposed to the accurate 21.893 slots.

λ	$\gamma=.0$	$\gamma=.3$	$\gamma=.5$
.04	8.500	14.928	23.500
.06	12.250	21.893	34.750
.08	23.500	42.785	68.500

Table I
Mean packet delay results.

References

1. D. Bertsekas, R. Gallager, "Data Networks", Prentice Hall, 1987.
2. I. Rubin, "Access-Control Disciplines for Multi-Access Communication Channels: Reservation and TDMA Schemes", IEEE Transactions on Information Theory, Vol. IT-25, No. 5, Sept. 1979.
3. I. Rubin, "Message Delays in FDMA and TDMA Communication Channels", IEEE Transactions on Communications, Vol. COM-27, No. 5, May 1979.
4. F. Hillier, B. Jabbari, "Analysis of the Fixed-Assigned TDMA Technique with Finite Buffer Capacity", IEEE Transactions on Communications, Vol. COM-35, No. 7, July 1987.
5. I. Stavrakakis, "A Statistical Multiplexer for Packet Networks", IEEE Transactions on Communications, submitted.
6. H. Heffes, D. Lucantoni, "A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Analysis", IEEE Journal on Selected Areas in Communications, Vol. SAC-4, No. 6, Sept. 1986.
7. A. Viterbi, "Approximate Analysis of TIme Synchronous Packet Networks", IEEE Journal on Selected Areas in Communications, Vol. SAC-4, No. 6, Sept. 1986.

IOANNIS STAVRAKAKIS

He received his Diploma in electrical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, in 1983 and his 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 include multi-user communication theory, stochastic processes, queueing theory and system performance evaluation.