Queueing Analysis of a Class of Star-Interconnected Networks under Markov Modulated Output Process Modeling

Ioannis Stavrakakis, Member, IEEE

Abstract—In this paper, a Markov modulated process is developed for the exact characterization of the output process of a class of random access networks (C-RAN’s) and the mean packet delay induced by the central node of a star interconnecting scheme is calculated. The latter is obtained through an approximate analysis of an appropriate queueing system under Markov modulated arrivals.

It is shown that the ALOHA network is a C-RAN. ALOHA networks interconnected according to a star topology are studied and delay results are obtained. These results are compared with those obtained under Bernoulli approximation on the network output processes and simulations.

I. INTRODUCTION

Packet communication networks have been widely adopted as an efficient means of transferring information. Such networks extend from small local area networks to large systems of interconnected networks covering extended geographical areas, [1]. The development of local area networks (or, in general, single networks) has been the focal point of extended research over the last two decades, [1]. The performance evaluation of these networks has been facilitated by the adopted models on the packet generation mechanism. Independent and identically distributed (i.i.d.) processes are widely used to model the per user packet generation process of a small (finite) user population network over a fixed time interval (slot). In the case of a large (infinite) user population system, the Poisson (or batch Poisson) model is usually adopted for the cumulative packet generation process. The previous models are in accordance with the randomness and the unpredictability of the packet generating mechanism, which is captured by the memoryless nature of these processes.

The single communication network imposes severe limitations on the information exchange capabilities of the supported users. To enhance these capabilities, simple communication networks are interconnected, resulting in larger and more complex systems, [2], [10]. At the same time, new network components are created to support the interconnection of the involved networks. These new components form the backbone network.

The basic problem in analyzing systems of interconnected networks is that of the characterization of the output processes generated by the supported networks. The characterization of this process is of fundamental importance to the analysis of interconnecting schemes. It is the input process to the interconnecting system and affects considerably its operation. In the case of a random access network (RAN) the packet output process is the process of the successfully transmitted packets, if all the packets are assumed to be forwarded to the interconnecting mechanism. Another problem is that of determining how a random access protocol operates in the presence of a node that forwards exogeneous traffic coming from other networks. The latter problem can be avoided by assigning a separate channel to the exogeneous traffic. In this case, the operation of the protocol is not affected by the exogeneous traffic, but the problem of optimum allocation of the available resources (channels), arises. The latter issue has been discussed in [11] where the objective is to maximize the throughputs of the interconnected networks.

The output process of a RAN depends on the protocol that has been deployed. Although its input process may be assumed to be memoryless, the random access protocol introduces dependencies to the output process. The description of this process is a difficult task and only approximations based on special assumptions have been attempted. The output process of RAN’s is a highly dependent process and memoryless models are meaningless. Despite it, many researchers have adopted such models due to the tractability of the resulting models. In [12], the Bernoulli model for the output process of a CSMA/CD network is implied. In [13] the authors consider the output process of ALOHA and CSMA networks by making the assumptions of the heavy traffic conditions and the memoryless property. Memoryless output processes are also implied in [14] and [15] in the analysis of two-hop ALOHA and CSMA packet radio networks and in [16] in the case of the multi-hop extension. The packet interdeparture process of ALOHA and CSMA networks is derived in [17].

In previous work, [18], [19], we developed Markovian approximations on the output process of a certain class of RAN’s, the objective being to capture, to some extend, the dependencies introduced by the random access protocol. This class contains all RAN’s whose analysis utilizes the process of the renewal points induced by the operation of the deployed protocol. Most continuous and limited sensing algorithms fall
into this category. The performance of such interconnected RAN's was evaluated by incorporating the Markovian approximations on the output processes of the involved networks.

In this paper a Markov modulated process is proposed for the description of the packet output process generated by certain RAN's; we call such RAN's C-RAN's. This process is an i.i.d. process whose parameters depend on the transitions of an underlying Markov chain describing the evolution of the RAN (Section II). In the sequel, a queueing system with Markov modulated arrivals (according to the model presented in Section II) is analyzed and the induced mean packet delay is calculated (Section III). This queueing system is used to model the central node of a star topology of interconnected C-RAN's. Some practical systems where star topologies of interconnected RAN's appear are presented (Section IV). In Section V, ALOHA networks are considered. It is shown that the finite population ALOHA network is a C-RAN. Its output process is exactly described by incorporating the Markov modulated process (defined in Section II) and a system of star interconnected ALOHA networks is studied by applying the queuing results derived in Section III). Finally, the conclusions of this work appear in the last section.

II. DESCRIPTION OF THE OUTPUT PROCESS

Consider a slotted communication network. It is assumed that the length of a slot is equal to the time required for a packet transmission; packet transmissions can be attempted only at the beginning of the slots. Network users are usually assumed to generate packets according to an i.i.d. compound process. In most cases, it is possible to describe the operation of the network by incorporating an appropriate Markov chain embedded at the beginning or the end of the slots. Given the state of the Markov chain, it may be possible to completely determine (probabilistically) the packet departure process from the network. Networks for which the latter is possible will be called C-NET's.

A C-NET and its packet output process is precisely determined through the following definitions.

Definition 1: The output process of a slotted communication network is defined to be the binary discrete time process \( \{a_j\}_{j \geq 0} \) of the departing packets; \( a_j = 1 \) if a packet leaves the network at the \( j \)th slot and \( a_j = 0 \) otherwise.

Note that in the case of a contention-free network \( \{a_j\}_{j \geq 0} \) is the process of the channel status (or activity). In the case of a random access network (RAN), \( \{a_j\}_{j \geq 0} \) is the process of the successfully transmitted packets.

Definition 2: Define \( C \) to be the class of slotted networks which satisfy the following.

a) There exists a finite-state Markov chain \( \{z_i\}_{i \geq 0} \) embedded at the slot boundaries which describes the evolution of the network. Let \( S = \{x_0, x_1, \ldots, x_M\} \) be the state space of \( \{z_i\}_{i \geq 0} \).

b) For any state transition (say from \( x_i \) to \( x_k \)), there exists a stationary probabilistic mapping \( a(x_i, x_k) : S \times S \rightarrow \{0, 1\} \), which describes the channel activity in the slot over which the state transition takes place. Let \( a(x_i, x_k) = 1 \) with probability \( \phi(x_i, x_k) \) and \( a(x_i, x_k) = 0 \) with probability \( 1 - \phi(x_i, x_k) \).

Networks which belong in the class \( C \) will be called C-NET's. C-NET's whose operation is determined by a random access protocol will be called C-RAN's.

Definition 3: Following definitions 1 and 2 we define the output process of a C-NET to be the process

\[ \{a_j\}_{j \geq 0} = \{a_j(x_i, x_k)\}_{j \geq 0}. \]

That is, the output process is described as a Markov modulated Bernoulli process; the output process is a Bernoulli process whose intensity depends on the state transition of an underlying Markov chain. To the best of our knowledge, this is the first time that such a process is proposed for the description of the output process of a (random access) communication network.

III. ANALYSIS OF A SINGLE SERVER QUEUEING SYSTEM

In this section we study a single server queueing system with Markov modulated arrivals. This general queueing model will be used for the analysis of star interconnect C-NET's. We believe that the significance and the applicability of this queueing model is beyond the particular application presented in this paper.

A. General Case: Asymmetric System

The asymmetry of the system is due to the fact that although all arrival processes are described by the same model, at least one of their parameters is considered to be different for at least two such processes.

Consider the queueing system that is shown in Fig. 1. It consists of \( N \) input streams which feed a single server. The server has an infinite capacity buffer. The arrival processes \( \{a^i_j\}_{j \geq 0} \), \( i = 1, 2, \ldots, N \), are assumed to be synchronized discrete time processes; at most one arrival can occur in each input line per unit time. The time separation between successive possible arrival points is constant and equal to one. The first in–first out (FIFO) policy is adopted and the service time is assumed to be constant and equal to the distance separation between successive time instants. More than one arrivals (from different input streams) that occur at the same time instant are served in a randomly chosen order.

Let \( \{z^j\}_{j \geq 0} \) denote a discrete time ergodic Markov process associated with the \( j \)th input stream, with finite state space \( S' = \{x'_1, \ldots, x'_M\} \). Let also \( a^j \) be a stationary probabilistic mapping from the set \( S' \times S' \) into the set \( \{0, 1\} \) where 1 corresponds to an arrival and 0 to no arrival. Then, we define
the arrival process of the $i$th input stream to be
\[ a_i^{(1) j=0} = \{a_i^{(1)}(j-1), a_i^{(1)}(j)\}_{j \geq 0}. \]
From the description of the arrival process it is implied that successive arrivals from the same input stream are not independent, but they are governed by an underlying finite state Markov chain, $\{z_i^{(1)}\}_{j \geq 0}$. The arrival process can also be seen as the random reward associated with a state transition of a Markov chain, \([20]\). Given a transition from state $k$ to state $j$, the arrival process is described by the probabilistic mapping $a_i^{(1)}(k, j)$ where $a_i^{(1)}(k, j) = 1$ with probability $\phi_i^{(1)}(k, j)$ and $a_i^{(1)}(k, j) = 0$ with probability $1 - \phi_i^{(1)}(k, j)$. It is assumed that the underlying processes $\{z_i^{(1)}\}_{j \geq 0}, i = 1, 2, \ldots, N$, are mutually independent and thus the arrival processes $\{a_i^{(1)}\}_{j \geq 0}, i = 1, 2, \ldots, N$, are also independent.

Previous work on similar queueing systems can be found in \([21]\)–\([25]\) (and the references cited there). All previous models differ significantly from the one presented here. In \([21]\), the authors assume a single arrival line and a two-state Markov modulate Poisson arrival process. In \([22]\), the author considers a single input line and arrivals that depend on an underlying two-state Markov chain. In \([23]\)–\([25]\), $N$ input lines are considered. In \([23]\) it is assumed that the packet arrival process of each of the identical input lines depends on an underlying two state Markov chain (active/inactive). In \([24]\) it is assumed that the per line packet arrival process is a first-order Markov chain and at most one packet arrival is possible. A closed-form solution for the mean packet delay has been derived for the latter case. In \([25]\), a closed-form expression for the mean packet delay in the case of Bernoulli per line arrival can be found. The systems presented in \([23]\)–\([25]\) (and some special cases of the system in \([22]\)), are special cases of the general system considered in this paper.

Let $x_i^{(k)}$ and $p_i^{(k)}(k, j), k, j \in S^2$, denote the steady state and the transition probabilities of the ergodic Markov chain, $\{z_i^{(1)}\}_{j \geq 0}, i = 1, 2, \ldots, N$. Let $p_i^{(j)}(k)$ denote the joint probability that there are $j$ packets in the system at the $n$th time instant or slot boundary (arrivals at that slot boundary are included) and the states of the Markov chains are $x_i^{(k)}, y_i^{(j)}, \ldots, y_i^{(n)}$ where $y_i^{(j)} = (y_i^{(1)}, y_i^{(2)}, \ldots, y_i^{(n)})$. The vector $y_i^{(j)}$ describes the state of a new ergodic Markov chain that is generated by $N$ independent Markov chains described before, with steady state probabilities $\pi_i(\bar{x})$, transition probabilities $p_i(\bar{x}, \bar{y})$ and state space $\bar{S}$ given by
\[
\pi_i(\bar{x}) = \prod_{i=1}^{N} \pi_i^{(1)}(x_i^{(k)}), \quad p_i(\bar{x}, \bar{y}) = \prod_{i=1}^{N} p_i^{(1)}(x_i^{(k)}, y_i^{(j)}),
\]
\[ \bar{S} = S^1 \times S^2 \times \cdots \times S^N. \]
(1)

The operation of the system can be described by an $N + 1$ dimensional Markov chain imbedded at the slot boundaries, with state space $T = (0, 1, 2, \ldots) \times \bar{S}$. Let $w_{\bar{x}y}$ where
\[
w_{\bar{x}y} = \sum_{i=1}^{N} a_i^{(1)}(x_i^{(k)}, y_i^{(j)}),
\]
be a random variable describing the cumulative packet arrivals which result from a state transition from $\bar{x}$ to $\bar{y}$; let $g_{\bar{x}y}(\nu), 0 \leq \nu \leq h, \mu_{\bar{x}y}$ and $\sigma^2_{\bar{x}y}$ be the probability distribution, the mean and the variance of $w_{\bar{x}y}$, respectively. The state probabilities of the $N + 1$ dimensional Markov chain are given by the following equations:
\[
p_i^{(n)}(j; \bar{y}) = \sum_{x \in S} \sum_{\nu=0}^{N} p_i^{(n-1)}(j + 1 - \nu, x)p(x, \bar{y})g_{\bar{x}y}(\nu)
\]
\[ j \geq N + 1 \]
\[
p_i^{(n)}(j; \bar{y}) = \sum_{x \in S} \sum_{k=1}^{j+1} p_i^{(n-1)}(k; x)p(x, \bar{y})g_{\bar{x}y}(j + 1 - k)
\]
\[ + \sum_{x \in S} p_i^{(n-1)}(0; x)p(x, \bar{y})g_{\bar{x}y}(j), \quad 0 \leq j \leq N \]
(3a)

(3b)

where $\bar{x}$ is the state of the $N$-dimensional Markov chain at time instant $n - 1$. Equation (3a) is easily understood by noting that the $N + 1$ dimensional Markov chain is in state $(j; \bar{y})$ at the $n$th slot boundary if it were in state $(j + 1 - \nu; \bar{x})$ in the previous slot boundary and a transition from $\bar{x}$ to $\bar{y}$ took place resulting in the arrival of $\nu$ packets. Equation (3b) is explained in a similar way. There are totally $M^1 \times M^2 \times \cdots \times M^N$ equations given by (3) for a fixed $j$ and all $\bar{y} \in \bar{S}$ where $M^i$ is the cardinality of $S^i, i = 1, 2, \ldots, N$.

Ergodicity of the Markov chains associated with the input streams implies the ergodicity of the arrival processes $\{a_i^{(1)}\}_{j \geq 0}, i = 1, 2, \ldots, N$. The latter together with the ergodicity condition for the total average input traffic $\lambda$
\[
\lambda = \sum_{x \in S} \sum_{\bar{y} \in \bar{S}} \mu_{\bar{x}y}p_{\bar{x}y}(\bar{x}, \bar{y})\pi(\bar{x}) < 1
\]
(4)

imply that the Markov chain described in (3) is ergodic. The steady-state (equilibrium) probabilities can be derived by considering the limit of the equations in (3) as $n$ approaches infinity, and obtain similar equations for the steady-state probabilities. By considering the generating function of these probabilities, manipulating the resulting equations, differentiating with respect to $z$ and setting $z = 1$, we obtain the following system of linear equations (see Appendix A).
\[
P_i^{(1)}(\bar{x}) = \sum_{x \in S} P_i^{(1)}(x, \bar{y})p(x, \bar{y}) + \sum_{x \in S} (\mu_{\bar{x}y} - 1)p(x, \bar{y})\pi(\bar{x})
\]
\[ + \sum_{x \in S} p(0; x)p(x, \bar{y})p(\pi(\bar{x}), \bar{y} \in \bar{S}.
\]
(5)

In general, the exact calculation of the boundary joint probability $p(0; \bar{x})$ is complicated. We use the following expression to estimate its value
\[
p(0; \bar{x}) = \frac{\sum_{x \in S} p(x, \bar{x})\pi(\bar{x})q(\bar{x}, \bar{y})}{\sum_{x \in S} \pi(\bar{x})p(\bar{x}, \bar{y})q(\bar{x}, \bar{y})}
\]
(6)

where $p_0 = 1 - \lambda$ is the probability that there is no customer in the system and $q_0(\bar{x}, \bar{y})$ is the probability that the state transition from $\bar{x}$ to $\bar{x}$ does not occur in no customer arrival; the latter
probability is easily obtained from the probabilistic mappings in the independent streams and it is given by

$$q_0(\bar{x}, \bar{x}) = \prod_{i=1}^{N} \left(1 - \phi \left( z^i, x^i \right) \right). \quad (7)$$

Notice that the calculation of the above boundary probability is the only point of approximation in this work. This is the reason for which we derive simulation results in the application presented in Section V. The exact calculation of the boundary probability is easy when for each state $\bar{x}$ there is only one state $\bar{x}_0$ such that a transition from $\bar{x}_0$ to $\bar{x}$ may result in no packet arrival; all other transitions into state $\bar{x}$ will result in at least one packet arrival. In this case, the boundary probability is easily derived to be equal to

$$p(0; \bar{x}) = p_0 q_0(\bar{x}_0, \bar{x}). \quad (8)$$

The $M^1 \times \cdots \times M^N$ linear equations with respect to $\bar{y} \in \bar{S}$ that appear in (5) are linearly dependent. This is the case when the equations are derived from the state transition description of a Markov chain. By manipulating the original equations and using L'Hospital's rule we obtain an additional linear equation with respect to $P'(1; \bar{y})$, $\bar{y} \in \bar{S}$, (see Appendix B) which is linearly independent from those in (5) and it is given by

$$\sum_{\bar{x} \in \bar{S}} \left[ 2(\mu_{\bar{x}} - 1)P'(1; \bar{x}) + 2(\mu_{\bar{x}} - 1)p(0; \bar{x}) + 2 + \sigma^2_{\bar{x}} + (\mu_{\bar{x}})^2 - 3\mu_{\bar{x}} \right] \pi(\bar{x}) = 0 \quad (9)$$

where

$$\mu_{\bar{x}} = E_{\bar{y}}[\mu_{\bar{y}}], \quad \sigma^2_{\bar{x}} = E_{\bar{y}}[\sigma^2_{\bar{y}}]. \quad (10)$$

By solving the $M^1 \times \cdots \times M^N$ dimensional linear system of equations which consists of (9) and any $M^1 \times \cdots \times M^N - 1$ equations taken from (5), we can compute $P'(1; \bar{x})$, $\bar{x} \in \bar{S}$. Then, the average number of packets in the system $Q$ can be computed by adding all the solutions. The average time $D$ that a packet spends in the system can be obtained by using Little’s formula as the ratio $Q/\lambda$. This is the mean delay of an arbitrary packet arriving at any of the input streams.

Consider the special case in which the per-stream arrival process is Bernoulli. The underlying Markov chain has one state and (5) and (9) become

$$p(0) = 1 - \mu \quad (5')$$

and

$$2(\mu - 1)P'(1) + 2(\mu - 1)p(0) + \left[ 2 + \sigma^2 + \mu^2 - 3\mu \right] = 0 \quad (9')$$

where

$$\mu = \sum_{i=1}^{N} \lambda_i, \quad \sigma^2 = \sum_{i=1}^{N} \lambda_i(1 - \lambda_i) \quad (11)$$

where $\lambda_i$ is the rate of the $i$th network. From (5') and (9') we get the following equation with respect to $P'(1)$

$$P'(1) = Q_B = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j + \mu(1 - \mu)}{(1 - \mu)} \quad (12)$$

where $Q_B$ is the average number of packets in the system. The mean packet delay, $D_B$ is given by $Q_B/\mu$, which is a known result. [25].

B. Special Case: Symmetric System

Let us now assume that the queueing system is symmetric, that is, the parameters of all input processes are identical. Let $M$ be the cardinality of each of the involved one-dimensional Markov chains. As expected, the number of equations which need to be solved for the calculation of the mean delay in the queueing system is reduced significantly. This can be easily seen by observing that the unknown quantities in (5), $P'(1; \bar{x})$, are the same with those corresponding to a permutation of $\bar{x}$.

If $\bar{v}(\bar{x}) = (v_1(\bar{x}), v_2(\bar{x}), \ldots, v_M(\bar{x}))$ is an $M$-dimensional vector with $v_i(\bar{x}), i = 1, 2, \ldots, M$, denoting the number of input processes at state $S_i$, then each vector $\bar{v}(\bar{x})$ with the constraint $\sum_{i=1}^{M} v_i(\bar{x}) = N$, represents a class of equivalent states $\bar{x}$. The number of equivalent states $\bar{x}$ in a class $\bar{v}(\bar{x})$ is given by

$$c(\bar{x}) = \binom{N}{v_1(\bar{x}), v_2(\bar{x}), \ldots, v_M(\bar{x})} \quad (13)$$

where $v_i(\bar{x})$, is the number of input processes in state $S_i$, as determined by $\bar{x}, i = 1, 2, \ldots, M$ [see [26, pp. 20]].

Let $F$ be the set of representative states $\bar{x}$ of the symmetric system (i.e., no two states $\bar{x} \in F$ belong to the same class of equivalent states); let $v(\bar{x})$ be the class of the equivalent to $\bar{x}$ states. For each $\bar{x}, \bar{y} \in F$, (5) and (9) can be written as follows:

$$P'(1; \bar{y}) = \sum_{\bar{x} \in F} \left\{ \sum_{\bar{x} \in V(\bar{y})} p(\bar{x}, \bar{y}) \right\} P'(1; \bar{x})$$

$$+ \sum_{\bar{x} \in S} \left[ (\mu_{\bar{x}} - 1) \right] p(0; \bar{x}) + 2 + \sigma^2 + \mu^2 - 3\mu \pi(\bar{x}) = 0 \quad (5a)$$

By solving the above equations with respect to $P'(1; \bar{x})$, $\bar{x} \in F$, we obtain the average number of customers in the
queueing system under input state $\bar{x}_{o}$, for each $\bar{x}_{o} \in F$. Then the average delay in the queueing system can be obtained from

$$D_{o} = \frac{\sum_{\bar{x}_{o} \in F} P^*(1; \bar{x}_{o})c(\bar{x}_{o})}{\lambda}$$

(14)

where $\lambda$ is given by (4). Depending on the number of input streams, the reduced number of equations $K$ in (5a) is easily computed. This number is given by

$$K = \frac{(M + N - 1)!}{N!(M - 1)!}$$

(15)

which is the number of ways of partitioning $N$ things into $M$ groups [26]. From (15) it turns out that significant reduction in the number of equations can be achieved, especially for small values of $M$, in the case of symmetric inputs. Partially symmetric inputs will also result in a reduction of the number of equations.

IV. STAR TOPOLOGIES OF INTERCONNECTED NETWORKS

The objective in this section is the presentation of some practical systems of interconnected RAN's where a star interconnecting scheme seems to be both meaningful and efficient.

Consider a system of two packet radio slotted communication networks which operate on virtually the same (or neighboring) areas (Fig. 2). In each network a random access protocol is employed for the allocation of the common channel to the network users. The users of each RAN can be either static or mobile. A central node, properly located, receives and retransmits the successfully transmitted packets coming from the corresponding network. Each RAN is assigned two channels, an uplink and a downlink. Thus, four different channels, to avoid interference, would be needed for the support of the two RAN's.

An important observation at this point is that the downlinks will be idle most of the time, due to the throughput limitations of the random access protocols. One way to save the bandwidth would be to divide unequally the bandwidth between the uplink and the downlink. This approach has been suggested in [11]. Another approach would be to use a common downlink for both RAN's. Comparison of the two approaches is beyond the scope of this paper. For the implementation of the latter scheme a common central node, receiving packets from both RAN's would be required. This structure saves one channel and offers an efficient solution to the problem of the interconnection of the involved networks. In fact, the latter is provided almost for free. At the same time a queueing system is formed at the common central node, due to the possibility of having simultaneous arrivals from both RAN's.

The system of the two RAN's described before can be extended to one which accommodates more than two RAN's, provided that the cumulative arrival rate to the central node is less than one packet per slot. The latter can be the case when the throughput of the RAN's is sufficiently small, or when a portion of the packet traffic arriving at the central node is not retransmitted because its destination is outside the interconnected RAN's.

Another system of practical interest is the one in which a buffer receives packets coming from (isolated) RAN's not to be retransmitted within the generating (or the interconnected) network, but to be forwarded to a distant destination via an appropriate transmission line (Fig. 3).

Clearly, the above are simple examples of star topologies of interconnected networks. The queueing system formulated in the central node of these topologies is hard to analyze. In the next section it is illustrated how a star topology of interconnected ALOHA RAN's can be analyzed by incorporating the Markov modulated model on their output process and using the analysis of the queueing system presented in Section III.

V. STAR-INTERCONNECTED ALOHA NETWORKS

In this section we analyze a star topology of interconnected ALOHA networks. At first, the output process of a slotted, single buffer, finite user population ALOHA network, [1], [14], [17], [27], [28], [29], is described. From the description turns out that this network is a $C$-RAN; we call it a $C$-ALOHA network. Then, the results from the analysis of the queueing system, presented in the previous section, are used for the calculation of the mean packet delay in the central node of a star topology of interconnected $C$-ALOHA networks.

Let $M$ be the number of users of the MUCN. A user can be either active (if its buffer is nonempty) or inactive (if its buffer is empty). An active user can be either a backlogged one (if its buffer was nonempty at the beginning of the current slot) or a new one (if its buffer was empty at the beginning of the current slot). The per user packet generation process is assumed to be Bernoulli with per slot probability of packet arrival $\lambda$. The
single buffer assumption implies that new packets which find the corresponding buffer full are discarded.

Two policies may be considered, the delayed first transmission (DFT) and the immediate first transmission (IFT).
Under the DFT policy, new and backlogged users (at the end of the last slot) transmit at the beginning of the current slot with probability $p$. Under the IFT policy, the backlogged users transmit at the beginning of a slot with probability $p$ and the new users transmit with probability 1.

Let us assume that the length of a slot equals one. We define the $j$th slot to be the time interval $(j, j + 1]$. Let $z_j$ be the number of active users (under the DFT policy) or the number of backlogged users (under the IFT policy) at the end of the $j$th slot. Under the IFT policy we assume that the new arrivals over a slot appear at the beginning of this slot, [17], [1]. It is easy to see that $(z_j)_{j \geq 0}$ is a Markov chain under both policies with state space $S = \{0, 1, 2, \ldots, M\}$.

The transition probabilities of this Markov chain have been derived for the analysis at these ALOHA protocols, [16], [29], [1], and are given in Appendix C. This Markov chain is finite and irreducible so it is ergodic, for all arrival rates, [1], [20]. Let $P$ denote the state transition probability matrix. The stationary distribution $\pi = (\pi(0), \pi(1), \ldots, \pi(M))$, where $\pi(k) = \lim_{j \to \infty} \Pr(z_j = k)$, is simply obtained by solving the system

$$H = \Pi P$$

Since $p(k, j) = 0$ for $j < k - 1$ the system can be solved recursively, [29].

Suppose that the Markov chain $(z_j)_{j \geq 0}$ moves from state $k$ at time $j - 1$ to state $i$ at time $j$. Let $a(k, i)$ be a binary random variable that describes the channel activity over the $j$th slot; $a(k, i)$ equals 1 if a successful packet transmission took place in the $j$th slot and is 0 otherwise; $a(k, i)$ is a Bernoulli random variable which is completely described by states $k$ and $i$ and the policy under consideration. More specifically, the expressions for the transition probabilities lead to the following:

$$a(k, i) = \begin{cases} 1 & \text{with probability } \phi(k, i) \\ 0 & \text{with probability } 1 - \phi(k, i) \end{cases}$$

where

a) under DFT policy

$$\phi(k, i) = \begin{cases} 0 & \text{if } i < k - 1 \\ \frac{\sigma_{k-1}}{\sigma_{k-1} + \tau_k} & \text{if } k - 1 \leq i \leq M \end{cases}$$

b) under IFT policy

$$\phi(k, i) = \begin{cases} 0 & \text{if } i < k - 1 \text{ or } i = k + 1 \text{ or } M \geq i \geq k + 2 \\ 1 & \text{if } i = k - 1 \\ \frac{\sigma_k}{\sigma_k + \tau_k} & \text{if } i = k \end{cases}$$

The involved quantities are defined in Appendix C.

The above system description, in view of Definitions 1–3, implies that the ALOHA network considered above is a C-RAN; its packet output process is completely described by the process $\{a(z_{j-1}, z_j)\}_{j \geq 0}$.

Consider $N$ C-ALOHA networks interconnected according to a star topology. Each of these networks is assumed to support $M$ users. The star topology may model a practical network interconnecting scheme, as discussed in Section IV. The central node of this topology is assumed to have the characteristics of the single server described in Section III. All networks are synchronized and have identical slot lengths. A packet departure from a network occurs at the end of a slot involved in a successful packet transmission and is declared as an arrival to the central node at the beginning of the next slot. Clearly, the output process of the $i$th C-ALOHA network, $\{a_i^j\}_{j \geq 0} = \{a_i^j(z_{j-1}, z_j)\}_{j \geq 0}$, is the arrival process of the $i$th input stream, according to the terminology of the previous section; superscript $i$ denotes quantities associated with the $i$th network.

The mean time that a packet spends in the central node of the star interconnecting topology is given by the solution of the linear equations given by (5) and (9) [or by (5a) and (9a) for the symmetric case] where the following expressions are used for certain quantities (assuming a transition from $x_i^i$ to $y_i^i$):

$$\mu_{xy} = E \left( \sum_{i=1}^{N} a_i^i(x_i^i, y_i^i) \right) = \sum_{i=1}^{N} E\{a_i^i(x_i^i, y_i^i)\}$$

$$= \sum_{i=1}^{N} \phi_i^i(x_i^i, y_i^i)$$

$$\phi_i^i(x_i^i) = \sum_{j=0}^{M} \phi_i^j(x_j^i, y_i^j)p_i^j(x_j^i, y_i^j)$$

$$\sigma_{xy}^2 = \sum_{i=1}^{N} (1 - \phi_i^i(x_i^i)) \phi_i^i(x_i^i)$$

$$\mu_{xy} = E_0[\mu_{xy}] = \sum_{i=0}^{N} \sum_{j=0}^{M} \phi_i^j(x_j^i, y_i^j)p_i^j(x_j^i, y_i^j)$$

$$= \sum_{i=1}^{N} \phi_i^i(x_i^i)$$

$p_i^j(x_j^i, y_i^j), \pi(x_i^i)$ are calculated from (16) by using (C.1) and (C.2); $\phi_i^j(x_j^i, y_i^j)$ are given by (18a), (18b).

Numerical results for the mean delay in the central node of $N = 2$ and $N = 3$ ALOHA networks interconnected according to a star topology and operating under the DFT policy have been obtained by solving the equations in (5) and (9) [or (5a) and (9a) for the symmetric case]. In Table I, delay results are shown for the simple case of $M = 2$ users per network and for $N = 2$ and $N = 3$ interconnected networks. Similar results are shown for the case of $M = 10$ users per network and for $N = 2$ in Table II and $N = 3$ in Table III. Results are shown under the Bernoulli approximation on the network processes as well. Both results are compared with simulations. The results show that both the Bernoulli approximation and the developed exact model on the network output processes perform satisfactorily under light traffic. When the traffic increases, the developed model clearly outperforms the Bernoulli approximation.
TABLE I
RESULTS FOR THE MEAN PACKET DELAY IN THE CENTRAL NODE OF A STAR TOPOLOGY OF N = 2 AND N = 3 INTERCONNECTED ALOHA NETWORKS UNDER DFT POLICY. λ_in is the per Network Input Rate, λ_out is the per Network Output Rate, p is the Packet Transmission Probability, D_net is the Network Induced Delay, D_q,n is the Queuing Delay Under the Developed Model, D_q,n−− is the Queuing Delay From the Simulations and D_q,n−− is the Queuing Delay Under the Bernoulli Model.

<table>
<thead>
<tr>
<th>M = 2 users per network</th>
<th>λ_in</th>
<th>λ_out</th>
<th>p</th>
<th>D_net</th>
<th>D_q,2</th>
<th>D_q,2−−</th>
<th>D_q,3</th>
<th>D_q,3−−</th>
<th>D_q,3−−−</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.098</td>
<td>0.86</td>
<td>1.40</td>
<td>1.06</td>
<td>1.06</td>
<td>1.14</td>
<td>1.14</td>
<td>1.14</td>
<td>1.14</td>
</tr>
<tr>
<td>0.20</td>
<td>0.188</td>
<td>0.82</td>
<td>1.64</td>
<td>1.14</td>
<td>1.15</td>
<td>1.15</td>
<td>1.40</td>
<td>1.43</td>
<td>1.43</td>
</tr>
<tr>
<td>0.30</td>
<td>0.266</td>
<td>0.78</td>
<td>1.85</td>
<td>1.25</td>
<td>1.27</td>
<td>1.28</td>
<td>2.16</td>
<td>2.20</td>
<td>2.31</td>
</tr>
<tr>
<td>0.35</td>
<td>0.300</td>
<td>0.77</td>
<td>1.96</td>
<td>1.32</td>
<td>1.35</td>
<td>1.37</td>
<td>3.56</td>
<td>3.62</td>
<td>3.96</td>
</tr>
<tr>
<td>0.40</td>
<td>0.330</td>
<td>0.75</td>
<td>2.06</td>
<td>1.41</td>
<td>1.44</td>
<td>1.49</td>
<td>28.78</td>
<td>29.39</td>
<td>34.00</td>
</tr>
<tr>
<td>0.50</td>
<td>0.381</td>
<td>0.73</td>
<td>2.25</td>
<td>1.66</td>
<td>1.70</td>
<td>1.80</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>0.60</td>
<td>0.419</td>
<td>0.70</td>
<td>2.44</td>
<td>2.07</td>
<td>2.12</td>
<td>2.29</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>0.70</td>
<td>0.447</td>
<td>0.68</td>
<td>2.61</td>
<td>2.29</td>
<td>2.37</td>
<td>2.51</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>0.80</td>
<td>0.480</td>
<td>0.64</td>
<td>2.79</td>
<td>2.98</td>
<td>4.00</td>
<td>4.45</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
</tbody>
</table>

TABLE II
RESULTS FOR THE MEAN PACKET DELAY IN THE CENTRAL NODE OF A STAR TOPOLOGY OF N = 2 INTERCONNECTED ALOHA NETWORKS UNDER DFT POLICY. λ_in is the per Network Input Rate, λ_out is the per Network Output Rate, p is the Packet Transmission Probability, D_net is the Network Induced Delay, D_q,2 is the Queuing Delay Under the Developed Model, D_q,2−− is the Queuing Delay From the Simulations and D_q,2−−− is the Queuing Delay Under the Bernoulli Model.

<table>
<thead>
<tr>
<th>M = 10 users per network</th>
<th>λ_in</th>
<th>λ_out</th>
<th>p</th>
<th>D_net</th>
<th>D_q,2</th>
<th>D_q,2−−</th>
<th>D_q,3</th>
<th>D_q,3−−</th>
<th>D_q,3−−−</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.099</td>
<td>0.51</td>
<td>2.40</td>
<td>1.07</td>
<td>1.06</td>
<td>1.06</td>
<td>1.5</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td>0.190</td>
<td>0.41</td>
<td>3.70</td>
<td>1.16</td>
<td>1.16</td>
<td>1.16</td>
<td>1.5</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>0.30</td>
<td>0.265</td>
<td>0.33</td>
<td>5.41</td>
<td>1.27</td>
<td>1.28</td>
<td>1.28</td>
<td>1.5</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>0.40</td>
<td>0.320</td>
<td>0.29</td>
<td>7.51</td>
<td>1.40</td>
<td>1.42</td>
<td>1.42</td>
<td>1.5</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>0.350</td>
<td>0.24</td>
<td>9.62</td>
<td>1.54</td>
<td>1.55</td>
<td>1.55</td>
<td>1.5</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>0.60</td>
<td>0.366</td>
<td>0.21</td>
<td>11.03</td>
<td>1.65</td>
<td>1.65</td>
<td>1.65</td>
<td>1.5</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>0.70</td>
<td>0.375</td>
<td>0.18</td>
<td>13.36</td>
<td>1.75</td>
<td>1.75</td>
<td>1.75</td>
<td>1.5</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>0.80</td>
<td>0.380</td>
<td>0.17</td>
<td>14.80</td>
<td>1.74</td>
<td>1.74</td>
<td>1.74</td>
<td>1.5</td>
<td>1.5</td>
<td></td>
</tr>
</tbody>
</table>

Notice that as long as the per network packet generation rate is less than 0.4 (so that the packet rejection probability is small), the induced queuing delay is less than half a packet length in the case of N = 2 networks. This was expected since the total output rate from both networks is less than 0.65, well below the capacity limit of the server which is 1. In the case of N = 3 networks, the total packet departure rate from all networks can be as high as the capacity limit of the server. Under such rates the queuing delay introduced by the interconnecting topology can be arbitrarily high as the capacity limit of the server is reached.

VI. CONCLUSION

In this paper, we have proposed a Markov modulated process for the description of the dependent output process of communication networks and we have analyzed a multi-input single server queue that models the central node of a star topology of interconnected networks.

The Markov modulated process has been shown to be capable of describing exactly the packet output process of a class of communication networks, the C-NET's. This class consists of all networks whose operation can be described by a finite state Markov chain and the output process can be obtained by a well defined probabilistic mapping from the state space of the Markov chain into the (departure/no-departure) set. The proposed model is exact and, unlike the i.i.d. approximation, captures the dependencies present in the network output process. It has been shown that the single buffer, finite population ALOHA network is a C-NET. By using the results of the analysis of the single server queue, a star topology of interconnected C-ALOHA networks has been approximately analyzed, under the Markov modulated model for the output processes. The results have shown that the proposed model outperforms the i.i.d. approximation on the network output process.

APPENDIX A

In this section we derive the linear equations that are given by (5). We write the steady state probabilities (under ergodicity) by considering the limit of (3) as n → ∞ and obtain

\[ p(j; y) = \sum_{x \in S} \sum_{\nu=0}^{N} p(j+1 - \nu; x)p(x, y)g_{xy}(\nu), \]

\[ j \geq N + 1 \]

\[ p(j; y) = \sum_{x \in S} \sum_{k=0}^{j+1} p(j-k, x)p(x, y)g_{xy}(j+1 - k) \]

\[ + \sum_{x \in S} p(0, x)p(x, y)g_{xy}(j), \]

\[ 0 \leq j \leq N. \]
### Table III

Results for the mean packet delay in the central node of a star topology of $N = 3$ interconnected ALOHA networks under DFT policy. $\lambda_{in}$ is the per network input rate, $\lambda_{out}$ is the per network output rate, $D_{net}$ is the network induced delay, $D_{i,j}$ is the queuing delay under the developed model, $D_{i,j-1}$ is the queuing delay from the simulations, and $D_{i,j}$ is the queuing delay under the Bernoulli model.

<table>
<thead>
<tr>
<th>$\lambda_{in}$</th>
<th>$\lambda_{out}$</th>
<th>$p$</th>
<th>$D_{net}$</th>
<th>$D_{i,j}$</th>
<th>$D_{i,j-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.099</td>
<td>0.51</td>
<td>2.40</td>
<td>1.15</td>
<td>1.14</td>
</tr>
<tr>
<td>0.20</td>
<td>0.190</td>
<td>0.41</td>
<td>3.70</td>
<td>1.48</td>
<td>1.45</td>
</tr>
<tr>
<td>0.30</td>
<td>0.268</td>
<td>0.33</td>
<td>5.41</td>
<td>2.31</td>
<td>2.27</td>
</tr>
<tr>
<td>0.40</td>
<td>0.320</td>
<td>0.29</td>
<td>7.51</td>
<td>7.14</td>
<td>6.80</td>
</tr>
<tr>
<td>0.43</td>
<td>0.329</td>
<td>0.26</td>
<td>8.12</td>
<td>22.64</td>
<td>19.79</td>
</tr>
</tbody>
</table>

If $P(z; \bar{y})$ is the z-transform of the joint probability distribution that there are $j$ packets in the system and the Markov chain is in state $\bar{y}$, defined by

$$P(z; \bar{y}) = \sum_{j=0}^{\infty} p(j; \bar{y})z^j$$

then we obtain

$$P(z; \bar{y}) = \sum_{j=0}^{N} p(j; \bar{y})z^j + \sum_{j=N+1}^{\infty} \sum_{\nu=0}^{N} p(j+1-\nu; \bar{x})$$

$$\cdot p(\bar{x}, \bar{y})z^j g_{\bar{y}y}(\nu)$$

$$= \sum_{j=0}^{N} p(j; \bar{y})z^j + \sum_{j=N+1}^{\infty} \sum_{\nu=0}^{N} p(j+1-\nu; \bar{x})$$

$$\cdot p(\bar{x}, \bar{y})z^j g_{\bar{y}y}(\nu)$$

$$= \sum_{j=0}^{N} p(j; \bar{y})z^j + \sum_{\nu=0}^{N} \sum_{k=N+2-\nu}^{\infty} g_{\bar{y}y}(\nu)$$

$$\cdot \left[ \sum_{k=N+2-\nu}^{\infty} p(k; \bar{x})p(\bar{x}, \bar{y})z^k g_{\bar{y}y}(\nu) \right]$$

$$= \sum_{j=0}^{N} p(j; \bar{y})z^j + \sum_{\nu=0}^{N} z^{\nu-1}$$

$$\cdot \left[ P(z; \bar{x}) - \sum_{k=0}^{N+1-\nu} p(k; \bar{x})z^k \right]$$

$$\cdot p(\bar{x}, \bar{y})g_{\bar{y}y}(\nu)$$

$$= \sum_{j=0}^{N} p(j; \bar{y})z^j + \sum_{\nu=0}^{N} z^{\nu-1}$$

$$\cdot P(z; \bar{x})g_{\bar{y}y}(\nu)$$

$$- \sum_{\nu=0}^{N} \sum_{\bar{x} \in S} z^{\nu-1}p(0; \bar{x})p(\bar{x}, \bar{y})g_{\bar{y}y}(\nu)$$

$$= \sum_{j=0}^{N} p(j; \bar{y})z^j + \sum_{\nu=0}^{N} z^{\nu-1}$$

$$\cdot P(z; \bar{x})g_{\bar{y}y}(\nu)$$

$$- \sum_{\nu=0}^{N} \sum_{\bar{x} \in S} z^{\nu-1}p(0; \bar{x})p(\bar{x}, \bar{y})g_{\bar{y}y}(\nu)$$

By differentiating the above equations, setting $z = 1$ and carrying out the sum with respect to $\nu$ we obtain the equations in (5).

### Appendix B

To derive the additional equation that is shown in (9), we differentiate (A1) and then add all the resulting equations. Since

$$P'(z) = \sum_{\bar{x} \in S} P'(z; \bar{x}) = \sum_{\nu=0}^{N} \sum_{\bar{x} \in S} P'(z; \bar{x})g_{\bar{x}y}(\nu)$$

$$g_{\bar{x}y}(\nu) = \sum_{\nu=0}^{N} g_{\bar{x}y}(\nu)$$ (B0)
we obtain

\[ P'(z) = \sum_{\nu=0}^{N} \sum_{\gamma \in S} \left[ (\nu - 1)z^{\nu-2}[P(z; \gamma) + (z - 1)p(0; \gamma)] 
\right. \\
\left. + z^{\nu-1}[P'(z; \gamma) + p(0; \gamma)] \right] g_{\gamma}(\nu). \]

By adding

\[ zP'(z) - z \sum_{\nu=0}^{N} \sum_{\gamma \in S} P'(z; \gamma) g_{\gamma}(\nu) \]

to the second part of above equation and rearranging terms we obtain that

\[ P'(z) = \frac{A(z)}{1 - z} \] \hspace{1cm} \text{(B1)}

where

\[ A(z) = \sum_{\nu=0}^{N} \sum_{\gamma \in S} \left[ (\nu - 1)z^{\nu-2}[P(z; \gamma) + (z - 1)p(0; \gamma)] 
\right. \\
\left. + z^{\nu-1}[P'(z; \gamma) + p(0; \gamma)] \right] - zP'(z; \gamma) g_{\gamma}(\nu). \]

Since, \( P'(z) \) is the average number of packets in the system, which is finite if (4) holds, and since \( 1 - z = 0 \) for \( z = 1 \), we compute \( P'(1) \) by applying L'Hospital's rule to (B1); thus

\[ P'(z) \bigg|_{z=1} = \frac{dA(z)/dz}{d(1 - z)/dz} \bigg|_{z=1} = -dA(z)/dz \bigg|_{z=1}. \] \hspace{1cm} \text{(B2)}

From condition (B2) and by using (B0) we obtain (9).

APPENDIX C

In this Appendix we give the transition probabilities describing the Markov chain of the ALOHA networks described in Section V.

a) Under DFT Policy:

\[ p(k, j) = \begin{cases} 0 & j < k - 1 \\ \tau_{kj} + \sigma_{kj} & k - 1 \leq j \leq M \end{cases} \] \hspace{1cm} \text{(C.a)}

where \( \tau_{kj} \) denotes the probability that there is no successful packet transmission when \( k \) users compete for the channel and that \( j - k \) from the inactive users are activated; similarly, \( \sigma_{kj} \) denotes the probability that there is a successful packet transmission when \( k \) users compete for the channel and that \( j - k + 1 \) of the inactive users are activated. These probabilities are given below where \( b_k(1) \) denotes the probability of a successful transmission, i.e., a single transmission, when \( k \) users compete for the channel

\[ \tau_{kj} = [1 - b_k(1)] \binom{M - k}{j - k} (1 - \lambda)^{M-j}, \]

\[ \sigma_{kj} = b_k(1) \binom{M - k + 1}{j - k + 1} (1 - \lambda)^{M-j}, \]

and \( b_k(0) = (1 - p)^k, \) \( b_k(1) = kp(1 - p)^{k-1} \) and \( \binom{m}{k} = 0 \) for \( k < 0. \)

b) Under IFT Policy:

\[ p(k, j) = \begin{cases} 0 & j < k - 1 \\ b_k(1)(1 - \lambda)^{M-k} & j = k - 1 \\ \tau_{kj} + \sigma_{kj} & j = k \\ (M - j)(1 - \lambda)^{M-k-1}(1 - b_k(0)) & j = k + 1 \\ \binom{M-k}{j-k} (1 - \lambda)^{M-j} & j \geq k + 2 \end{cases} \]

where

\[ \tau_{kj} = \left[ 1 - b_k(1) \right] (1 - \lambda)^{M-k} \]

and \( \sigma_{kj} = (M - k)\lambda(1 - \lambda)^{M-k-1}b_k(0). \)

These equations are explained as those under the DFT policy.

REFERENCES


Ioannis Stavrakakis (S’85–M’89) was born in Athens, Greece, in 1960. He received the Diploma in electrical engineering from the Aristotlean 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 and the Technical Chamber of Greece.