AN ANALYSIS APPROACH TO MULTI LEVEL NETWORKING

Ioannis Stavrakakis

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

Abstract

The major problem in the analysis of large packet communication networks is related to the difficulty in characterizing the involved packet traffics. In this paper, a general description of the packet processes generated by network components is presented. Based on this description, an analysis approach to multi level networking is developed. The packet traffic generated by a lower level (which is the input traffic to an upper level) is described as a generalized Bernoulli process whose intensity depends on the state of that level. This characterization is shown to be exact, or a well performing approximate one, for many practical systems. The general approach is illustrated through the analysis of a specific multi level packet communication network.

I. Introduction

In this paper, a network component is defined as a network subsystem which generates packets. Such components can be a local area network, a network switch or a repeater, a statistical multiplexer, a link that carries traffic from another large network, or any line carrying packetized information of virtually any kind. We define the First Stage Component (**FSC**) to be a network component whose packet input process is determined by simple first-born traffic. The simple first-born traffic is defined as a traffic which is not modulated by any network component. This type of traffic is modeled as an independent process.

In a large packet communication network (defined here as a network with more than one components), there is always a mechanism which provides for the backbone network that supports the interconnection of the involved components. Clearly, (some of) the packet processes in this system are the output processes from the supported components. We define the Second Stage Component (**SSC**) to be a system whose input process is (at least partially) determined by the output processes of other network components.

The performance evaluation of a large network requires the analysis of all FSCs and SSCs. The analysis of FSCs has received much attention in the past and many analytical techniques have been developed for that purpose. The nature of the models for the simple first-born traffic has played an important role in the development of these techniques. The analytical tractability provided by the Bernoulli / Poisson models for the simple first-born traffic has tempted many researchers to adopt these models for the description of the input processes to SSCs as well, [1]-[3]. Such models are usually inaccurate since, although the input process to a FSC might be considered as memoryless, the FSC introduces dependencies to its output process.

In this paper, an alternative to the i.i.d. (independent and identically distributed) characterization of the packet (output) processes generated by the components of a communication network, is presented. The proposed model is exact and it is presented in the next section. The output process is described as a generalized Bernoulli process whose intensity depends on the state of an underlying Markov chain describing the operation of the component. The cardinality of the state space of this Markov chain may be almost arbitrarily large. To provide for a numerically tractable solution in the later case, the exact model is appropriately modified. For this purpose, a meaningful approximate model on the output process, based on the dominant states of the underlying Markov chain, is introduced. The approximate model is described by introducing the concept of the α%-exact model.

II. Models for the output processes.

II.1 The exact model.

The description of the dependent packet process generated by a network component (FSC or SSC) is essential to the analysis of a SSC since the input processes to the latter are determined by the packet

processes generated by the feeding network components. The proposed exact model on the output (or packet departure) process of a network component is described through the following definitions. The discussion in this paper is confined to discrete time network components, as formulated by slotted packet communication systems.

Definition 1 The packet generation (output) process of a network component is defined to be the discrete time process of the departing packets, $\{a_j\}_{j\geq 0}$; $a_j=\rho$, $0\leq\rho<\infty$, if ρ packets leave the component at the jth time instant.

Definition 2 Assume that the network component satisfies the following:

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

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

Then, the output process of the component 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. Notice that the process $\{a_j\}_{j\geq 0}$, as given by (1), describes exactly the output process of a network component, provided that the conditions in Definition 2 are satisfied.

II.2 The α%-exact model.

To define the α%-exact model on the dependent process generated by a network component consider two processes $\{z_j\}_{j\geq 0}$ and $\{a_j(z_j)\}_j\geq 0$, as described in Definition 2, associated with the component. We define the new Markov chain $\{z_j'\}_{j\geq 0}$ with state space $S'=\{x_1', \cdots ,x_{M'}'\}$ for some $M'<M$. The states $x_1', \cdots ,x_{M'-1}'$ of the new Markov chain are identical to the $M'-1$ dominant states of the original Markov chain; $x_{M'}'$ is the union of the remaining states of the original Markov chain. Dominant states are the states with the largest probability mass. The steady state and state transition probabilities of the new Markov chain are obtained by appropriate averaging of those of the original Markov chain. We define $\{a_j'\}_{j\geq 0}$ to be a process similar to $\{a_j\}_{j\geq 0}$ with a corresponding underlying Markov chain $\{z_j'\}_{j\geq 0}$ and probabilistic mapping $a'(z_j'): S'\rightarrow Z_0$ obtained by appropriate averaging of that of the original output process.

Clearly, the process $\{a_j'\}_{j\geq 0}$ is an approximation on the true packet output process of the network component described by $\{a_j\}_{j\geq 0}$. The approximation is introduced by the reduction of the state space of the original underlying Markov process $\{z_j\}_{j\geq 0}$, which describes completely the operation of the component. In view of the construction of the approximate process $\{a_j'\}$, it is reasonable to expect that as M' increases, $\{a_j'\}_{j\geq 0}$ approaches $\{a_j\}_{j\geq 0}$; that is, the approximate model approaches the exact one. When $M'=M$, the approximate model coincides with the exact one. It is also reasonable to expect that the larger the total probability mass of the (unchanged) $M'-1$ states the better the approximation achieved by the reduced state space process. In view of the previous observations, we provide the following definition-measure of the approximation on the exact process $\{a_j\}_{j\geq 0}$.

Definition 3 The process $\{a_j'\}_{j\geq 0}$ defined as before is called α% -exact if the cardinality M' of the corresponding approximate underlying Markov chain $\{z_j'\}_{j\geq 0}$ (as defined before) is such that,

$$\sum_{1\leq k\leq M'-1} \pi'(x_k')\geq\alpha \quad \text{and} \quad \sum_{1\leq k\leq M'-2} \pi'(x_k')<\alpha$$

That is, the total probability mass of the unchanged states of the original Markov chain $\{z_j\}_{j\geq 0}$ is at least α and the total probability mass of the $M'-2$ dominant states is less than α.

Assuming ergodicity of all processes involved, Definition 3 can be roughly interpreted in the following way: An α%-exact output process $\{a_j'\}_{j\geq 0}$ is based on the true packet generating mechanism (state of the system) α% of the time; $(100-\alpha)$% of the time, the output process is based on an average packet generating mechanism. An average packet generating mechanism is the only one assumed present when $\{a_j\}_{j\geq 0}$ is approximated by a (generalized) Bernoulli process. The latter case corresponds to merging all states of the true underlying Markov chain into a single one ($M'=1$). An average number of outputs is generated under the latter model throughout the time horizon, independently of the true state of the underlying packet generating mechanism. In view of Definition 3, the exact model on the output process $\{a_j\}_{j\geq 0}$ corresponds to $\alpha=100$. $100-\alpha$ can also be seen as a measure of the smoothing on the output process introduced by the merging of the states of the true packet generating mechanism.

From the definition of the α%-exact process and the construction of $\{z_j'\}_{j\geq 0}$ turns out that the cardinality of the state space of $\{z_j'\}_{j\geq 0}$ (i.e. M') increases with α. Thus, it is reasonable to expect that the

larger the value of α the better the approximation on the packet output process. Assuming that this is generally true, there is a trade off between the degree of the accuracy of the approximating process and the introduced complexity in its description, as measured by M'. The real concern at this point is not about the complexity in the description of the approximating (or exact) packet output process itself, but it is on the tractability of the analysis of other components of a large network, whose input processes are described by the proposed models.

III. The output process of some network components

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 [4] it has been found that network packet traffic is bursty. As a result, a first order Markov model could be adopted in 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, [4], 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). The parameters of the packet output process of a network component which generates bursty traffic can be determined from the packet rate $\pi(1)$ and the burstiness coefficient γ.

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 $\sigma(i) = \Pr\{\text{message consists of i packets}\}$, $1 \le i \le K$. 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. 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 $\{z_j\}_{j\ge 0}$ and $\{a_j\}_{j\ge 0}$ defined in the previous section. If z_j is the number of packets in the node at the end of the jth slot, then it can be easily shown that $\{z_j\}_{j\ge 0}$ is a Markov chain with state space $S=\{0,1,2,\cdots,K\}$. The nonzero transition probabilities are given by

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

$$p(0,0)=p(1,0)=1-\mu\ \ ,\ \ p(k,k-1)=1 \text{ for } 2\le k\le K$$

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

$$\phi_1(k) = 1\ ,\ 1\le k\le K\quad ,\quad \phi_1(0) = 0$$

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

A Bernoulli approximate model on the output process of the node would have intensity $1-\pi(0)$. 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

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

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

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

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 $g(k)$, $0\le k\le K$, 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 $\{z_j\}_{j\ge 0}$ and $\{a_j\}_{j\ge 0}$. If z_j is the number of packets in the node at the end of the jth slot, then $\{z_j\}_{j\ge 0}$ is a Markov chain with state space $S = \{0,1,2,\ldots,M\}$. The transition probabilities are given by (assume $g(k)=0$ for $k>M$)

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

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

and the probabilistic mapping is given as in Example 2. The Bernoulli and the Markov approximations on the resulting packet output traffic can be determined as in the previous example.

IV. Performance analysis of a large network

In this section, the multi-component network

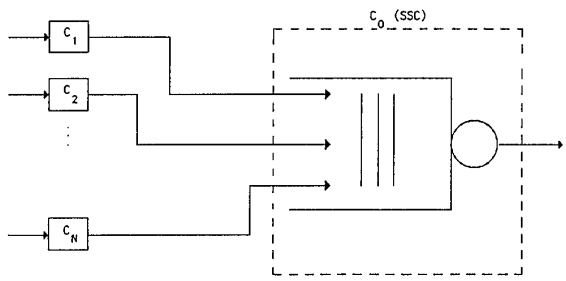

Figure 1.

shown in Fig. 1 is analyzed and the mean packet delay in C_0 is calculated. The SSC of this network is a common queueing system which has been analyzed under various input processes; the single server is assumed to have a buffer of infinite capacity. The arrival processes $\{a_j^i\}_{j\geq 0}, i = 1,2, \cdots ,N$, are assumed to be synchronized discrete time processes, and at most one arrival can occur in each input line per unit time. The time separation between successive potential packet arrival instants is constant (slot). The first in-first out (FIFO) policy is adopted and the service time is assumed to be constant and equal to one slot. More that one arrivals that occur at the same time instant are served in a randomly chosen order.

When each of the output process of the components $C_1, C_2,...,C_N$ is a Bernoulli or a first-order Markov process, the mean packet delay in C_0 is given by a closed form expression (see [4], [5]). When the output processes of the components $C_1, C_2, \cdots ,C_N$ are given by (1) (see section II), then the mean packet delay in C_0 is obtained from the solution of $M^1 x M^2 x \cdots M^N$, linear equations, [5], where M^i is the cardinality of the underlying Markov chain of the ith component (or by a smaller number of equations, in symmetric cases).

In the sequel, we derive some results for the mean packet delay in C_0 under various network components $C_1,...,C_N$ and under exact or approximate descriptions of their packet output processes (Fig.1). These results lead to some conclusions regarding some commonly made oversimplifying assumptions on the packet processes generated by network components. N=3 network components (other than C_0) are considered in the following cases.

Case 1:

Let C_1 be the network component described in Example 1 with parameters $\pi(1) = .1$ and $\gamma = .3$. Let C_2 be the network component described in Example 2 with parameters K=5, $\sigma(0) = 0$, $\sigma(1) = .1$, $\sigma(2) = .3$, $\sigma(3) = .3$, $\sigma(4) = .2$, $\sigma(5) = .1$ and $\mu = .1$, which result in a packet output rate of .244 packets per slot. Finally, let C_3 be the network component described in Example 3 with parameters M=50 (buffer size), $r=g(0) = .8$ (probability of no message arrival in a slot), $g(1)=.1(1-r)$,$g(2)=.3(1-r)$, $g(3)=.3(1-r)$ $g(4)=.2(1-r)$ and $g(5)= .1(1-r)$, where $g(k)$, $k=1,...,5$, is the probability that a message consists of k packets. The output processes of components C_1 and C_2 are exactly described as mentioned in section III. For the description of the output process of component C_3, both, the exact (which involves an underlying Markov chain with a state space of cardinality 51) and the a%-exact approximate (which involves an underlying Markov chain of cardinality less than 51) models, are adopted (see Section II).

For the network described above, the mean packet delay induced by the SSC C_0 is shown in Table I. Exact results, obtained by incorporating the exact models on the input processes to C_0, are shown. Results for various degrees of approximation, α, of the output process of C_3 are shown, as well. For a certain value of α, the range of the dominant states is also shown. It can be easily seen that, e.g., the 97%-exact model (resulting in a Markov chain of cardinality 10) gives mean delay results in C_0 very close to the accurate ones obtained by solving 5 times more linear equations. This observation indicates that a significant reduction in the numerical complexity of the problem can be achieved by incorporating an approximate model based on the dominant states of a component, at the expense of an insignificant deviation from the accurate results. Other possible approximations on the true output process, which also simplify the analysis, are the Bernoulli and the Markov ones. Results under these approximations are shown in Table I. As it can be easily concluded the Bernoulli model completely fails to approximate the accurate result, while the Markov model is clearly inferior to the α%-exact model for sufficiently large α. Some more conclusions about the performance of the a%-exact model are drawn in the next case.

Case 2:

Let C_3 be as in Case 1 and C_1 and C_2 be network components as described in Example 1. Let $\pi(1) = .17$ and $\gamma=.3$ for each of the components C_1 and C_2. The delay results in C_0 under various models on the output processes of C_1, C_2, and C_3 are given in Table II, for $\gamma=.3$, and in Table III for $\gamma=.0$; the latter value of γ corresponds to Bernoulli traffic.

For a very small value of α only the most significant state will be considered, while the rest of them will be merged into a single state. This situation is reflected for $\alpha=1$ where the state 0 (the most significant one) is selected. Let 1′ denote the new

Model	Delay
100%(0-50)	14.655
Bernoulli	4.174
Markov	13.427
85%(0-5)	13.720
90%(0-6)	13.959
95%(0-8)	14.295
97%(0-9)	14.388
99%(0-12)	14.515

Table I.

Model	Delay
100%(0-50)	12.884
Bernoulli	4.060
Markov	10.604
1%(0-0)	10.604
50%(0-1)	10.171
80%(0-4)	11.595
90%(0-6)	12.387
95%(0-8)	12.705
97%(0-9)	12.793
98%(0-10)	12.850

Table II

Model	Delay
100%(0-50)	11.404
Bernoulli	4.060
Markov	9.140
1%(0-0)	9.140
50%(0-1)	8.689
80%(0-4)	10.113
90%(0-6)	10.905
95%(0-8)	11.224
97%(0-9)	11.311
98%(0-10)	11.368

Table III

state generated by merging all other states 1-50. Since, for this particular component, state 0 results in no packet generation and state 1'(or states 1-50) always results in one packet generation, the 1%-exact model coincides with the first order Markov approximation! Thus, the delay results for both models are the same (10.604). Although one would expect monotone improvement of the accuracy of the obtained delay results as the number of states merged into a single state decreases, this turns out not to be the case. Although the approximation on the output process itself is refined, the effect of it on the improvement of the accuracy of the delay result is not well predicted due to the complexity of the queueing process and the significance of the probability mass which is concentrated in the new compound state. As this probability mass decreases, the delay results improve in a monotone fashion. In the example considered here, the α%-exact model gives monotonically improving delay results (which are better than those under the Markov model) when the compound state concentrates less than 20% of the probability mass.

References

1. H. Takagi, L. Kleinrock, "Output Processes in Contention Packet Broadcasting Systems", IEEE Transactions on Communications, Vol. COM-33, No. 11, Nov. 1985.
2. F.A. Tobagi, "Analysis of a Two-Hop Centralized Packet Radio Network-Part I: Slotted ALOHA", IEEE Transactions on Communications, Vol. COM-28, No. 2, Feb. 1980.
3. H. Takagi, L. Kleinrock, "Throughput-Delay Characteristics of Some Slotted ALOHA Multihop Packet Radio Network", IEEE Transactions in Communications, Vol. COM-33, No. 11, Nov. 1985.
4. A. Viterbi, "Approximate Analysis of Time-Synchronous Packet Networks", IEEE Journal in Selected Areas in Communications, Vol. SAC-4, No. 6, Sept. 1986.
5. I. Stavrakakis, "Interconnected Packet Communication Networks", Symbosium on Command and Control, Washington, D.C., June 26-28, 1989 (also submitted for publication).