Lateness Probability of a Retransmission Scheme for Error Control on a Two-State Markov Channel

Michele Zorzi, Senior Member, IEEE, and Ramesh R. Rao, Senior Member, IEEE

Abstract—In this paper, we study the performance of a simple retransmission-based error-control strategy for delay-constrained data communications over a bursty channel. Correlated errors are modeled as a two-state Markov process. A retransmission algorithm is used to correct errors, and the probability that a packet is not successfully delivered within \(D\) slots of its arrival is computed. In the presence of a smoothing buffer at the receiver, this is the probability that the jitter experienced by a packet is too large to be absorbed by the buffer itself. The cases of zero and nonzero roundtrip delay are studied separately, as are the conditions of perfect and imperfect feedback.

Our results relate the achievable quality-of-service and the amount of traffic that can be served to the packet-error process parameters, which in turn are induced by the physical layer specifications. This relationship between the traffic and the channel parameters can be useful in making admission control decisions and in assessing the effect of physical layer design on the performance of higher layer protocols.

Index Terms—ARQ, error control, error correlation, Markov model, queueing system.

I. INTRODUCTION

RETRANSMISSION-based error-control techniques have been used and studied for some time now. Nonetheless, the upsurge of interest in packet-switched multimedia communications has renewed interest in these techniques. In addition to the short propagation delays of high-speed radio links, three other reasons stand out.

Automatic repeat request (ARQ) techniques provide a way to trade off delay jitter for residual errors. In contrast, forward error correction (FEC) schemes typically allow for the tradeoff between data rate (redundancy) and residual errors. The jitter/error tradeoff is at the heart of managing packet-switched streaming services such as voice and video over wide area networks. In such an environment, delay jitter is endemic and its management is a crucial aspect of the service design.

A second aspect of ARQ implementations that is germane to the design of quality-of-service (QoS) cognizant networks is the relative ease of implementing traffic-sensitive service policies. For example, assuming the use of a discard-all-aged-packets policy, the transmitter, upon receipt of a negative acknowledgment, can decide on a per connection basis whether or not to retransmit the packet. This is in contrast to FEC schemes for unequal error protection that require the decoder to recognize the various streams and process them differently—a task that may require a more computationally intensive and energy-consuming receiving unit [1].

A third relevant property of ARQ is the ability to effect selective discard policies. For example, control policies used in wireline networks might allow congested interior nodes to selectively discard packets from certain layers of a layer-encoded image. Such policies are harder to carry over to a wireless channel where errors due to outages are not controlled. Nonetheless, the retransmission mechanism of ARQ allows us to regain this control. Assuming the use of a discard-all-aged-packets policy, one can once again selectively retransmit only some of the packets and allow others to expire.

These three factors reinforce the need to determine the relationship between lateness probabilities and packet error or dropping, which is the focus of this paper.

We consider a pair of users communicating over a channel which exhibits correlated packet errors. A retransmission algorithm is used to correct errors, and the probability that a packet is not successfully delivered within \(D\) slots of its arrival (lateness probability) is computed. In the presence of a smoothing buffer at the receiver, this is the probability that the jitter experienced by a packet is too large to be absorbed by the buffer itself. The cases of zero and nonzero roundtrip delay are studied separately, as are the conditions of perfect and imperfect feedback.

II. DELAY CONSTRAINED RETRANSMISSION: ZERO PROPAGATION DELAY

We consider a communications link with a relatively high data rate (of the order of 1 Mb/s) and subject to bursty errors. The time axis is slotted, and transmission occurs in blocks of \(N\) bits. We focus on a single pair of communicating users and for simplicity ignore all other users. Given the short range of communications in many third-generation systems, the propagation delay is initially assumed to be zero for simplicity in this section and is extended to nonzero values in Section III. In each slot, packets are independently generated at the transmitter according to some arrival process and queued in a buffer awaiting transmission.

A. Channel Model

Errors that occur on wireless channels are a function of specific propagation artifacts such as multipath and fading, which result in correlated errors. Models to deal with cor-
related errors have been proposed in the literature. Especially useful are Markov models whose special structure makes them analytically tractable [2]. In [3], Wang proposed a continuous Markov model for the fading envelope, and in [4], Wang and Moayeri considered a similar model with a finite number of states. The case with only two states dates back to the early work by Gilbert [5].

The traditional metric used for characterizing channel errors at the physical layer is the average bit-error rate (BER) and its computation is a well-known practice (e.g., see [6]). A little harder is the evaluation of the error process at the packet level, where the success of a packet depends jointly on the errors on the various symbols [7]. Also, from the perspective of higher layer applications, one is interested in block errors, since many if not all applications run on top of a link layer that exchanges blocks of data. For the Gilbert–Elliott channel model, it is possible to find exact expressions for the block-error rate [8], [9]. Also, the joint statistics of the block-error process, with and without interleaving and forward error correction, has been computed in [9]. The higher order characterization of the block-error process has a number of applications, one of which is to assess the accuracy of a Markov model at the block level. Based on the measure proposed in [3], it was found in [9] that the block-error process can be reasonably well modeled as Markov. Therefore, the error process at the packet level is modeled here as a first-order two-state Markov chain. A similar model was used to study the performance of Go-Back-N (GBN) ARQ in [10] and [11], where on the other hand the delay performance was not addressed, and no arrival process was considered.

Let 0 and 1 denote successful and erroneous transmission in a given slot, respectively, and let

$$P = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$$

be the transition matrix for the packet-error process. The average packet-error rate is then given by

$$\varepsilon = \frac{p_{01}}{p_{10} + p_{01}}.$$  

Although considering the error process appears to be the most direct way to model the channel for our purposes, it is useful to observe that the above model is equivalent to a Gilbert model in which the error probabilities in bad state and good state are 1 and 0, respectively. This model was called "simplified Gilbert model" in [8].

This Markovian model is easily extended to account for diversity. In the presence of fading, diversity techniques improve performance by using two (or more) suitably spaced antennas. If the antennas receive independently attenuated replicas of the same signal, the probability of failure is reduced. Many techniques have been studied to combine the signals received at the two antennas. Here, we will consider ideal switched diversity, where the system is able to recognize and select (on a packet-by-packet basis) the antenna with the larger signal-to-noise ratio.

Assume that a diversity of order two (i.e., two antennas) is employed and that the signals received at the two antennas fade independently. The channel can now assume three states: both channels are good (state 0), the two channels have different states (state 1), and both channels are bad (state 2). This three-state process is still Markov. If the transition matrix of each channel is as given in (1), then the transition matrix of the three-state process can be computed as

$$P^{(d)} = \begin{pmatrix} p_{00}^2 & 2p_{01}p_{00} & p_{01}^2 \\ p_{10}^2 & p_{11}^2 + p_{10}p_{01} & 2p_{10}p_{11} \\ p_{11}^2 & 2p_{10}p_{11} & p_{11}^2 \end{pmatrix}$$

with steady-state distribution given by

$$\pi^{(d)} = ((1-\varepsilon)^2, 2\varepsilon(1-\varepsilon), \varepsilon^2).$$  

In this three-state model, both states 0 and 1 correspond to a transmission success, whereas state 2 corresponds to a transmission failure. Extension of the two-state channel approach to this case is certainly feasible, since the previously considered transition diagram can be generalized to include this case.

Alternatively, we could map the above three-state channel to a new binary process, which is 0 if the channel is in state 0 or state 1 (i.e., at least one success) and is 1 if the channel is in state 2. It is possible to find the joint statistics of the channel state and, therefore, that of this new process. Now, applying the mutual information criterion of [3] and [12] to this latter process, computations show that a binary Markov approximation may be adequate in modeling the diversity channel, and therefore this model can be used instead of the more complex three-state model. Since in the diversity channel a failure occurs if both channels are in bad state, we have for the average packet-error rate $\varepsilon^{(d)} = \varepsilon^2$. Moreover, since $P^{[\text{state 1}| \text{state 1}]} = p_{11}^2$, we have for the inverse of the average burst length

$$p^{(d)}_{10} = 1 - p_{11}^2.$$  

With this choice of parameters, the binary Markov model can be used for the diversity case as well.

### B. No Feedback Errors

In this section, we identify the state variables and study their evolution. As a first approach, a perfect and instantaneous feedback channel is assumed so that at the end of each slot, the transmitter knows exactly whether or not the transmission was successful. Packets that were not successfully received are immediately retransmitted. Thus, during the good periods, one packet is transmitted in each slot, unless the queue is empty and there is no arrival in that slot. During the bad periods, no packets can be transmitted, and arrivals are queued up, waiting for the channel to become good again. Therefore, at the end of a bad period, there will usually be a backlog of packets that need to be transmitted.

The system described above can be represented by a Markov chain $X(n) = (\ell(n), i(n))$ with state space $\{(\ell, i), 0 \leq \ell \leq 1, 0 \leq i \leq N\}$.
For the packet arrival process, consider a Bernoulli model first. A packet arrival occurs with probability $\lambda$ at time $nT^+$, independent of $X(m)$ for all $m \leq n$. This packet is not counted in the state $X(n)$. If $X(n) = (0, \bar{0}), i > 0$, then a packet is transmitted and successfully received (departure). Even if $X(n) = (0, \bar{0})$, there is a departure if there is an arrival in slot $n$, since we assume that a packet arriving at time $nT^+$ can be transmitted in that slot. On the other hand, if $X(n) = (1, \bar{0}), i \geq 0$, then no departure can occur.

Thus, the transition probabilities from the states $(0, i + 1), (1, i), i \geq 0$ are as follows:

$$(0, i) \rightarrow (0, i + 1), (1, i)$$

$$(1 - \lambda)p_{11} + (1 - \lambda)p_{101}$$

whereas the only two possible transitions from state $(0, 0)$ are to $(0, 0)$ and $(1, 0)$, with probability $p_{00}$ and $p_{001}$, respectively.

By inspection of the transition diagram for the chain, one can write a set of balance equations which lead to the following steady-state distribution:

$$\pi(0, i) = \beta \pi(0, 0), \quad i \geq 0$$

$$\pi(0, 0) = \frac{(1 - \lambda)p_{10} + \lambda p_{00}}{p_{01}}$$

where

$$\beta = \frac{(1 - \lambda)p_{11} + \lambda p_{101}}{(1 - \lambda)p_{10} + \lambda p_{100}}$$

and $\pi(0, 0)$ is found from the normalization condition to be

$$\pi(0, 0) = \frac{[(1 - \lambda)p_{10} - \lambda p_{01}]p_{01}}{(1 - \lambda)[p_{10} + p_{101}][p_{10} + \lambda p_{00}]}$$

The steady-state distribution $\pi(\ell, i)$ is therefore given by

$$\pi(0, 0) = 1 - \frac{\varepsilon}{1 - \lambda}$$

$$\pi(0, i) = \frac{\lambda \varepsilon (1 - \beta) \beta^{i - 1}}{(1 - \lambda)}, \quad i \geq 1$$

$$\pi(1, i) = \varepsilon (1 - \beta) \beta^{i}, \quad i \geq 0$$

Note that the necessary condition for stability $\beta < 1$ leads to $\lambda < 1 - \varepsilon$, as expected.

Consider now the case in which the arrival process is general and may depend on the channel state. Following an approach similar to the one described in [14], we can solve for the joint steady-state distribution of the channel state and the queue size at the beginning of a slot $\pi(\ell, i)$. Let $a_{k\ell}$ be the probability that $k$ arrivals occur in a slot, given that the channel is in state $\ell$. and let

$$G_{\ell}(z) = \sum_{k=0}^{\infty} a_{k\ell} z^k$$

be the generating function of this distribution.

Taking into account the fact that arrivals occur at the beginning of a slot and can depart in the same slot (if good), we can write the following balance equations for the steady-state distribution:

$$(\pi(0, i), \pi(1, i))$$

$$= \left(\sum_{k=0}^{\infty} a_{0k} \pi(0, 0+1+k) + \delta(\ell) a_{00} \pi(0, 0)\right) P$$

where $\delta(0) = 1$ and $\delta(i) = 0$ for $i \neq 0$.

Taking $Z$-transforms, we obtain

$$P_0(z) = \sum_{k=0}^{\infty} \pi(\ell, k) z^k$$

is the partial generating function [14].

Solving for $P_0(z), P_1(z)$ yields the following expressions:

$$P_0(z) = \frac{\left(z - \delta(\ell) a_{00} \pi(0, 0)\right)p_{00} + p_{10}p_{01} - p_{10}p_{01} G_{01}(z)}{\Delta(z)}$$

$$P_1(z) = \frac{\left(z - \delta(\ell) a_{00} \pi(0, 0)\right)p_{01}}{\Delta(z)}$$

where

$$\Delta(z) = (z - p_{10}(G_0(z))(1 - p_{11}G_1(z)) - p_{10}p_{01} G_0(z)G_1(z))$$

Finally, we can find the value of $\pi(0, 0)$ by using the normalization condition

$$\lim_{z \to 1} [P_0(z) + P_1(z)] = 1.$$ (17)

C. Feedback Errors

When errors are present in the feedback direction, we are forced to increase the state space of the chain. The “channel state” is now a pair, whose first entry gives the state of the forward channel and the second entry gives the state of the feedback channel. If we take the binary value of such a combination, we can define the channel state $\ell$ as follows:

- $\ell = 0$: correct transmission, correct feedback (00)
- $\ell = 1$: correct transmission, erroneous feedback (01)
- $\ell = 2$: erroneous transmission, correct feedback (10)
- $\ell = 3$: erroneous transmission, erroneous feedback (11).

Let

$$Q = \begin{pmatrix} q_{00} & q_{01} \\ q_{10} & q_{11} \end{pmatrix}$$

be the transition matrix for the feedback channel, and let $P$ [as given in (1)] be that for the forward channel. Then, the transition matrix for the state of the pair of channels is given

1A more general analysis for processes of this kind is presented in [13].
by the following:
\[
\Theta = \begin{pmatrix}
\theta_{00} & \theta_{01} & \theta_{02} & \theta_{03} \\
\theta_{10} & \theta_{11} & \theta_{12} & \theta_{13} \\
\theta_{20} & \theta_{21} & \theta_{22} & \theta_{23} \\
\theta_{30} & \theta_{31} & \theta_{32} & \theta_{33}
\end{pmatrix} = P \otimes Q
\]

where $\otimes$ denotes the Kronecker product between matrices [2].

The approach used in the previous section continues to apply. A suitable set of balance equations can be written, and the steady-state distribution of the chain can be determined. In particular, one must partition the state space into two exhaustive and mutually exclusive subsets and equate the probability flows in the two directions across the boundary between the subsets, in order to obtain one such equation. Due to the regular structure of the transition diagram, it is possible to write the flow equations in the following manner:

\[
S(i) \triangleq \begin{pmatrix}
\pi(1, i) \\
\pi(2, i) \\
\pi(3, i) \\
\pi(0, i + 1)
\end{pmatrix}
\]

\[
AS(i) = BS(i - 1), \quad \text{for } i > 0
\]

\[
AS(0) = B_0 \pi(0, 0)
\]

where (details are in Appendix A)

\[
A = \begin{pmatrix}
-\lambda \rho & -\lambda & -\lambda & -\lambda \\
1 & -\lambda & -\lambda & -\lambda \\
1 & \lambda & -\lambda & -\lambda \\
1 & \lambda & \lambda & -\lambda & -\lambda \\
0 & 0 & \lambda & \lambda & -\lambda & -\lambda
\end{pmatrix}
\]

\[
B = \begin{pmatrix}
-\lambda \rho & -\lambda & -\lambda & -\lambda \\
-\lambda & \lambda & \lambda & -\lambda & -\lambda \\
-\lambda & \lambda & \lambda & -\lambda & -\lambda \\
0 & 0 & \lambda & \lambda & -\lambda & -\lambda
\end{pmatrix}
\]

\[
B_0 = \begin{pmatrix}
-\rho & -\rho & -\rho & -\rho \\
-\rho & -\rho & -\rho & -\rho \\
-\rho & -\rho & -\rho & -\rho
\end{pmatrix}
\]

\[
\mu e_j = 1 - \mu e_j = \sum_{m \leq j} \theta_{0m}
\]

and $\bar{\lambda} = 1 - \lambda$.

The above equations can be rewritten as

\[
S(i) = M^i M_0 \pi(0, 0), \quad i \geq 0
\]

where

\[
M = A^{-1} B \quad M_0 = A^{-1} B_0.
\]

Finally, the value of $\pi(0, 0)$ can be analytically computed from the normalization condition

\[
1 = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \pi(\ell, i)
\]

\[
= \pi(0, 0) \left( 1 + \sum_{i=0}^{\infty} 1 M^i M_0 \right)
\]

\[
= \pi(0, 0) (1 + 1 [1 - M^{-1} M_0])
\]

\[
= \pi(0, 0) (1 + 1 [1 - M^{-1} M_0])
\]

where $1$ is a row vector of all 1’s.

\[\text{D. Computation of the Lateness Probability}\]

In the previous sections, we derived the steady-state distribution $\pi(\ell, i)$. Let us consider now the probability that a packet is not transmitted within $D$ slots of its arrival. In a delay constrained system, this will be the probability that a packet exceeds the maximum tolerable delay.

Let $\phi_{ij}(k, n)$ be the probability that there are $k$ successful slots in $\{0, 1, \ldots, n-1\}$ and that the channel state is $j$ at time $n$, given that the channel state was $\ell$ at time 0. For simplicity, we consider in the following case the case of perfect feedback, although the same technique applies to the case of feedback errors as well. Then, the following recursion, which applies for $n \geq 0$, can be obtained using the technique of [15]

\[
\phi_{ij}(k, n) = \phi_{ij}(k-1, n-1) p_{ij} + \phi_{ij}(k, n-1) p_{ij} + \delta_{ij} \delta(l) \delta(n)
\]

\[\text{where it is understood that } \phi_{ij}(k, n) = 0 \text{ for negative values of either } k \text{ or } n, \delta_{ij} = 1 \text{ for } i = j \text{ and zero otherwise, and } \delta(n) = \delta_{n0}. \text{ In this case, the closed-form solution for the recursion is given by (31), shown at the bottom of the next page (see Appendix B for the derivation).}\]

Let us consider a packet arriving in slot $n$, and let $(\ell, i)$ be the state of the system at time $nT$. Then, the delay suffered by the packet is given by the time it takes to transmit $i + 1$ packets (i.e., to have $i + 1$ successes), given that the channel started in state $\ell$. In particular, the probability $P_L(\ell, i)$ that the packet is not transmitted within $D$ slots, conditioned on the channel being in state $(\ell, i)$ in the slot of its arrival, is given by

\[
P_L(\ell, i) = \sum_{k=0}^{i} \phi_{ij}(k, D)
\]

\[\text{where } \phi_{ij}(k, n) = \phi_{ij}(k, n) + \phi_{ij}(k, n). \text{ Note that } P_L(0, 0) = 0 \text{ for any } D > 0, \text{ since if a packet arrives with the system in state (0, 0), it is immediately (and successfully) transmitted. Also, } P_L(\ell, i) = 1 \text{ for } i \geq D, \text{ if first-input first-output queueing service is assumed.}\]

The steady-state distribution of $X(n)$ seen by an arrival in slot $n$ is the steady-state distribution, since arrivals are memoryless. The unconditional probability that a packet spends more than $D$ slots in the queue is therefore given by

\[
P_L = \sum_{i=0}^{\infty} [P_L(0, i) \pi(0, i) + P_L(1, i) \pi(1, i)]
\]

\[\text{It is shown in Appendix C that } P_L = \frac{\varepsilon}{1 - \lambda} \rho^{D-1}
\]

\[\text{where } \rho = 1 - \frac{1 - \lambda - \varepsilon}{(1 - \varepsilon)(1 - \lambda)} p_{00}.
\]

If a packet-dropping mechanism that discards a packet $D$ slots after its arrival is used, then the above probability gives a conservative estimate, since the evolution with dropping is always stochastically dominated by that without dropping. The performance of the system with dropping is considered in Section IV-E. The case with finite buffer can be studied similarly [16]. In the numerical results section, we plot $P_L$ and discuss its forms.
III. DELAY CONSTRAINED RETRANSMISSIONS: FUTURE PRODUCTION DELAY

Consider now the case when the roundtrip delay is \( m > 1 \) slots, i.e., the acknowledgment of a block transmitted in slot \( n \) is not received at the end of slot \( n + m - 1 \). To study this case, we need to specify the protocol used for retransmissions; we shall assume here that GBN [17] is used, and that there are no feedback errors. The extension to include feedback errors is tedious, although conceptually straightforward, and is not considered here. The system model is the same as used in the previous analysis, except that now \( m \) can be greater than one.

As long as correct transmissions occur, a packet is transmitted in each slot (provided that the queue is not empty). If a packet is unsuccessfully sent in slot \( n \), the transmitter will be notified of the error by the end of slot \( n + m - 1 \) and will retransmit that packet in slot \( n + m \). If this retransmission is still unsuccessful, another copy of the packet will be sent in slot \( n + 2m \) and so on until in slot \( n + jm \) the packet is eventually successfully received, and normal operation is resumed. Note that according to the GBN protocol rules, all packets transmitted in slots \( n + 1 \) through \( n + jm - 1 \) (except for the retransmissions of the packet under consideration) are ignored by the receiver (even if successful).

Therefore, when in the bad state, the protocol effectively samples the channel with period \( m \) slots. If the channel becomes bad when the queue is not empty, then this sampling starts in the first bad slot. On the other hand, if the queue is empty, then this sampling has a random offset with respect to the arrival of the bad period, depending on when the first arrival occurs. In fact, the status of this sampled version of the channel depends on the arrival process. For analytical convenience, we assume that this offset is not random by forcing the transmissions during a bad period to occur in slots which are separated from the first bad channel slot by an integer multiple of \( m \). This has no effect if the channel turns bad when the queue is not empty. On the other hand, if it happens when the queue is empty, then the first arrival is not transmitted immediately but has to wait for the next slot in which a transmission can occur. This approximation of the system operation considerably simplifies the analysis, and is expected to have a minor impact when \( m \) is small.

Let us therefore consider the following sampling rule for the channel process. If the \( i \)th slot of the sampled version of the process corresponds to slot \( n \) in the original sequence, then the \( (i+1) \)th slot of the sampled process corresponds to slot \( n+1 \) if slot \( n \) was good and to slot \( n+m \) if slot \( n \) was bad. This sampled version of the channel process is a Markov chain with transition matrix

\[
P' = \begin{pmatrix}
P_{00} & P_{01} \\
P_{10} & P_{11}
\end{pmatrix} = \begin{pmatrix}
0 & 0 \\
1 & 0
\end{pmatrix} P + \begin{pmatrix}
1 & 0 \\
0 & 0
\end{pmatrix} P
\]

i.e., transitions from state 1 (the bad state) occur on an \( m \)-slot basis, whereas transitions from state 0 (the good state) occur at every slot.

The steady-state joint distribution of the channel state and the queue size at the beginning of a slot in this sampled process \( \zeta(\ell, i) \) can be computed from the analysis in Section II. A good and a bad slot in the sampled process correspond to one and \( m \) slots in the actual time frame, respectively. Assuming Bernoulli arrivals, i.e., one or zero arrivals in each actual slot with probability \( \lambda \) and \( 1 - \lambda \), respectively, we have

\[
G_0(z) = 1 - \lambda + \lambda z \quad \text{and} \quad G_1(z) = G_0^n(z).
\]

The partial generating functions \( R_0(z) \) and \( R_1(z) \) are given by (14)–(16), where

\[
\lim_{z \to 1} [R_0(z) + R_1(z)] = \frac{\alpha \alpha \zeta(0, 0)(p_{00} + p_{01})}{(1 - \lambda)p_{00} - \lambda mp_{01}}
\]

so that

\[
\zeta(0, 0) = \frac{(1 - \lambda)p_{10} - \lambda mp_{01}}{(1 - \lambda)(p_{00} + p_{01})}.
\]

For \( m = 1 \) (zero propagation delay), (40) reduces to

\[
\zeta(0, 0) = \frac{(1 - \lambda)p_{10} - \lambda p_{01}}{(1 - \lambda)(p_{00} + p_{01})}
\]

equal to what was found in Section II.

The above equations give the partial generating functions of the queue size, which can be inverted to get \( \xi(\ell, i) \) for \( \ell = 0, 1 \) and for all \( i \geq 0 \). Note that \( \zeta(\ell, i) \) is the distribution
in a random slot, whereas the distribution seen by a random arrival can be found as
\[
\zeta'(0,i) = \frac{\zeta(0,i)}{\sum_{k=0}^{\infty} [\zeta(0,k) + m\zeta(1,k)]}
\]
(42)
\[
\zeta'(1,i) = \frac{m\zeta(1,i)}{\sum_{k=0}^{\infty} [\zeta(0,k) + m\zeta(1,k)]}
\]
(43)

A. Delay Analysis

Consider next the probability that a packet is not transmitted within \(D\) slots of its arrival. Let \(\phi_{ij}(k,n)\) be the joint probability that \(k\) slots in \([0, \ldots, n-1]\) (in the actual time scale) are successful and that the channel state in slot \(n\) is \(j\), given that it was \(i\) at time 0. The recursion (30) can now be rewritten as [15, Ch. 10.6]

\[
\phi_{0j}(k,n) = \phi_{1j}(k-1,n-1)p_{10} + \phi_{0j}(k-1,n-1)p_{00} + \delta_{0j}\delta(n)\delta(k)
\]
(44)
\[
\phi_{1j}(k,n) = \phi_{1j}(k,n-m)p_{11} + \phi_{0j}(k,n-m)p_{00} + \delta_{1j}\delta(k)\sum_{\ell=0}^{m-1}\delta(n-\ell)
\]

which accounts for the fact that transitions from state 1 take \(m\) actual slots.

Given these functions, we compute the conditional lateness probability \(P_L(\ell,i)\), i.e., the probability that a packet which finds the (sampled) system in state \((\ell,i)\) is not transmitted within \(D\) slots since its arrival.

If \(\ell = 0\), then we simply have
\[
P_L(0,i) = \sum_{j=0}^{i} \phi_{0j}(j,D)
\]
(45)
where \(\phi_{0j}(j,n) = \phi_{00}(j,n) + \phi_{1j}(j,n)\).

If \(\ell = 1\), then in the sampled model, we have a batch of \(n_m\) arrivals in a bad slot, where \(n_m\) has binomial distribution with parameters \(\lambda, m\). Let us assume that an arrival occurs in a bad slot, and let \(k = 0\) and \(k = 1, \ldots, m-1\) indicate the bad slot itself and the following \(m-1\) slots on the actual time scale (up to the one which immediately precedes the next transmission attempt), respectively. Then, the actual slot in which the arrival occurs can be any integer between 0 and \(m-1\), with equal probability \(1/m\). Let \(P_L(1,i,k)\) be the lateness probability obtained by further conditioning on the packet arriving in slot \(k\). Then, the arriving packet will find \(i + n_k\) packets in the queue, with \(n_k\) binomial with parameters \(\lambda, k\). Moreover, \(D\) slots from the packet arrival correspond to \(D+k\) slots from the reference slot. Therefore, given \(n_k\) arrivals, \(P_L(1,i,k)\) is the probability that no more than \(i + n_k\) packets are successfully transmitted within \(D+k\) slots, given that the first slot was bad. Thus, \(P_L(1,i,k)\) is given by
\[
P_L(1,i,k) = \sum_{n_k=0}^{k} \binom{k}{n_k} \lambda^{n_k} (1-\lambda)^{k-n_k} \sum_{j=0}^{i+n_k} \phi_{1j}(j,D+k)
\]
(46)
and
\[
P_L(1,i) = \sum_{k=0}^{m-1} \frac{1}{m} P_L(1,i,k),
\]
(47)

IV. Numerical Results

In this section, we present some numerical results based on (33) and (48) for the system without dropping on a Markov channel. The performance is given in terms of the lateness probability \(P_L\) versus the arrival rate \(\lambda\), the average packet-error rate \(\varepsilon\), the average length of a burst of packet errors \(1/p_{10}\), and the delay bound \(D\). The parameters of the packet-error process can be related to the physical channel characterization as explained in [9], [12], and [18].

A. Perfect Feedback

The dependence of the lateness probability \(P_L\) on the various system parameters is studied in Figs. 1–5. Unless otherwise noted, we will assume an arrival rate of \(\lambda = 0.9\) and an average burst length \(1/p_{10} = 5\) slots, which corresponds, for example, to a fade duration of 5 ms on a 1-Mb/s channel when 1000-bit packets are transmitted.

Fig. 1 shows \(P_L\) versus the average packet-error rate \(\varepsilon\) for various values of the delay \(D\). Both absence (solid lines) and presence (dotted lines) of ideal double selection diversity, various values of the maximum tolerable delay \(D\).

Finally, the lateness probability is given by
\[
P_L = \sum_{i=0}^{D-1} \left[ \zeta'(0,i) P_L(0,i) + \zeta'(1,i) P_L(1,i) \right] + \left( 1 - \sum_{i=0}^{D-1} \zeta'(0,i) + \zeta'(1,i) \right).
\]
(48)

In this section, we present some numerical results based on (33) and (48) for the system without dropping on a Markov channel. The performance is given in terms of the lateness probability \(P_L\) versus the arrival rate \(\lambda\), the average packet-error rate \(\varepsilon\), the average length of a burst of packet errors \(1/p_{10}\), and the delay bound \(D\). The parameters of the packet-error process can be related to the physical channel characterization as explained in [9], [12], and [18].
The dependence on the channel burstiness is studied in Fig. 4, where $P_L$ is plotted versus $1/p_{10}$ for $D = 50$ slots and various values of the average packet-error rate. It is seen that the delay performance of the ARQ scheme degrades as the average burst length increases. More insight into the effect of the channel burstiness on the performance is given in Fig. 5, in which the Markov parameters, instead of being independently selected, are computed based on a Gilbert error process at the bit level [9]. In this case, changing the channel burstiness at the bit level affects both the burstiness at the packet level and the packet-error rate. In the right portion of the graph, long packet-error bursts result in high lateness probability, whereas in the presence of very isolated errors, the packet-error rate is very high. It can be seen that there exists an "optimal" value of the channel burstiness and that this optimal value is relatively independent of the delay constraint. The behavior of the curves in Fig. 5 suggests that techniques aimed at modifying the burstiness of the error process (such as interleaving) could be used in this case to "set" the correlation of the error process to its optimal value.

From the above results, we note that for typical values of the error process parameters, in most cases, the probability that a packet suffers a delay $D$ or larger can be made sufficiently low by the use of diversity. For example, for $\varepsilon$ as large as 0.1 and for $D = 50$ slots (which is less than 100 ms for ATM cells over a 256-kb/s channel and even less for higher data speed, as expected in future systems), we have $P_L < 4 \times 10^{-6}$, which is only slightly higher than the probability of an undetected error guaranteed by a standard 32-bit cyclic redundancy check, commonly accepted by most applications.

The results presented in this paper can be applied to the performance evaluation of higher layer protocols in the presence of bursty packet losses, e.g., as happen over wireless channels. In this case, an important issue is the interaction of delay induced by the lower layers and protocol actions (e.g., timeouts) at the higher layers. Being able to quantify the delay distribution at the lower layers enables one to...
appropriately design the timing of the higher layer protocol, avoiding mismatch situation which can potentially cause large performance penalties [19].

B. Admission Control Curves

Consider (34), which gives the relationship among the following parameters: $\varepsilon, 1/p_{10}$ (i.e., the channel description), $D, P_L$ (QoS specification), and $\lambda$ (traffic capacity). Let us fix the values of the QoS parameters, corresponding to a given application requirement. Since $P_L$ is a monotonic function of all its parameters (as can be easily verified directly), the equation

$$P_L(\varepsilon, p_{10}, D, \lambda) = P^*_L$$  (49)

when all parameters but one are kept fixed and has a unique solution, if one exists.

We can use this observation to answer the following question: given some QoS specification $(D, P^*_L)$, what is the achievable traffic region, i.e., the maximum arrival rate which guarantees the specified lateness, as a function of the channel description? The result of this calculation is presented in a three-dimensional plot in Fig. 6, where $\lambda$ is plotted versus $\varepsilon$ and $1/p_{10}$ (on a logarithmic scale) for $D = 50$ and $P^*_L = 10^{-4}$. As expected, when the channel conditions are good, $\lambda \approx 1$ can be served, whereas as the channel conditions degrade (i.e., $\varepsilon$ or $1/p_{10}$ or both increase), $\lambda$ decreases until it eventually reaches zero. Note that in a certain region (identified by the plateau at $\lambda = 0$ in the plot), the lateness condition cannot be satisfied. This happens when

$$\varepsilon (1 - p_{10})^{D - 1} > P_L^*$$  (50)

i.e., even the probability that packets arriving to an empty queue are delayed more than $D$ slots exceeds the lateness requirement.

Another way to look at these results is presented in Fig. 7, where the results of Fig. 6 are shown with a contour representation. Curves represent constant values of $\lambda$, i.e., the intersections of the surface of Fig. 6 with horizontal planes. The bolder curve represents the relationship (50), which corresponds to $\lambda = 0$. Results of this sort are very useful in relating the achievable QoS and the amount of traffic that can be served to the packet-error process parameters, which in turn are induced by the physical layer specifications. From this point of view, we can see different physical layer design choices as points in the $\varepsilon, 1/p_{10}$ plane. These choices can be directly compared by using the constant-$\lambda$ contours.

The relationship between $\lambda$ and the channel parameters, which gives the equivalent capacity as a function of the packet-error process characterization, can be useful in making admission control decisions. These results are also useful in assessing the effect of physical layer design on the performance of higher layer protocols.

C. Effect of Feedback Errors

The results of Section IV-A assumed the availability of an error-free feedback channel. This is not always true, especially in mobile radio systems where the feedback channel is itself a wireless channel subject to fading.

As an example of the effect of feedback errors, results for the Markov channel without diversity are presented in Fig. 8, where an average burst length of 5 packets is assumed and $\lambda = 0.9$. Curves are plotted for various values of the average packet-error rate. It can be seen that the degradation introduced by feedback errors is not very significant for the cases shown, i.e., the presence of an unreliable feedback link does not have a catastrophic effect. In the presence of diversity, feedback errors introduce an even smaller degradation due to increased channel reliability.

From the results presented, we conclude that an unreliable feedback channel introduces some degradation. However, if the error rate on the feedback channel is adequately small, the impact is small, so that ARQ schemes still remain competitive for the environment studied. Moreover, we assumed that the error structure on the two channels was the same, whereas in reality, feedback information is usually more protected than the data, e.g., some form of coding could be used. Note, in fact, that the impact of using error-correction in the feedback
D. Effect of the Propagation Delay

The analysis developed in Section III allows us to study how the propagation delay affects the performance. As an example of the results obtained, Fig. 9 shows the lateness probability $P_L$ versus the value of the roundtrip $m$ (with $m = 1$ corresponding to instantaneous feedback) for various values of $\varepsilon$ and two values of the average burst length, namely $1/p_{10} = 3$ and 10 slots. It can be seen that the degradation due to an increased propagation delay is reasonably small for moderate values of the roundtrip delay. On the other hand, for large values of $m$, the results with instantaneous feedback may provide an overly optimistic estimate of the performance. Also, the performance degradation with increasing $m$ is almost negligible in the presence of long channel memory [$1/p_{10} = 10$ (dotted lines)], whereas it is more pronounced for faster channel variations.

E. Effect of Packet Dropping

For analytical convenience, in the above analysis, we considered the case in which packets exceeding their delay constraint are just delivered late, i.e., the delay bound is not “hard” and delayed delivery, although undesirable, can be accepted. On the other hand, there are instances in which the delay constraint must be met, and late packets cannot be used at the receiver (packetized voice or video are good examples). In this case, it is obviously better to drop at the transmitter those packets which have reached their delay constraint, since they are going to be discarded at the receiver anyway and, by staying in the queue, cause additional delay to packets that arrived later and still may have a chance to meet their own deadline. A packet-dropping mechanism can therefore be envisioned, whereby the lifetime of the packets in the transmission queue is monitored and packets whose delay constraint is violated are dropped. The performance in this case can be expected to be better than in the system without dropping, so that the benefit of using an ARQ error-control strategy is even more significant. Unfortunately, the analysis of the scheme with packet dropping is exceedingly more difficult than the one considered in this paper, and therefore, we resort to simulation in order to gain some understanding of its performance.

As an example, Fig. 10 gives results for the case without diversity. Two values of average error rate, i.e., $\varepsilon = 0.01$ and 0.0001, are considered, and the performance of the system without dropping (analytically derived) and with dropping (obtained by simulation) is shown. Simulation for the case without dropping is also incorporated. It can be seen that dropping late packets significantly enhances the performance, with a reduction of the lateness probability by about one order of magnitude. This result reinforces our conclusion that ARQ proves to be an effective means to implement error control in this environment. Fig. 11 also shows that the dropping mechanism tends to make the system more stable, as expected.

It is interesting to note that the curves with dropping in Fig. 10 are parallel to those without dropping, i.e., our simple analysis seems able to predict the rate at which the
performance improves as the delay constraint is relaxed. Thus, by using the slope of the analytical results along with a single simulation point, we can provide in this case approximate results for small values of the lateness probability, which would be unachievable by pure simulation. Further investigation is needed to verify the sensitivity of this behavior to the system parameters and to the traffic model. Proper understanding of this behavior would be very useful for the study of the performance of schemes in which late packet dropping is implemented.

V. CONCLUSION

In this paper, we have considered ARQ error control for a wireless channel characterized by small propagation delay and error bursts, in the presence of a delay constraint. It has been shown that a simple ARQ protocol can effectively support a high-traffic intensity, while guaranteeing a delay constraint with high probability. Additional benefits of the scheme are simple implementation of on-demand QoS and of priorities. Also, the scheme turns the cause of packet loss from the errors on the channel to congestion in the queue, allowing the implementation of selective packet dropping.

APPENDIX A

DERIVATION OF (23)–(25)

The state space of the chain is given by \{\{(\ell,i), \ell = 0, \cdots, 3, i \geq 0\}\}. Let the states be ordered in the following manner:

\[(0,0), (1,0), (2,0), (3,0), (0,1), (1,1), (2,1), (3,1), (0,2), (1,2), (2,2), (3,2), \cdots \]  

(51)

and let the term “stage \(n\)” denote the block of states with \(i = n\).

We can find a balance equation by subdividing the state space into two mutually exclusive subsets. In particular, we can associate one such partition to each state \((j,\bar{j})\) by putting in the first subset the states from \((0,0)\) through \((j,\bar{j})\) (included), and in the second subset, all states which follow \((j,\bar{j})\) in the ordering of (51). The corresponding balance equation is then obtained by equating the probability flows between the two subsets.

Consider the case \(i > 0\). A state \((\ell, i-1)\) in stage \(i-1 \) contributes probability flow in the forward direction (i.e., from left to right) only if there is a transition leading from \((\ell, i-1)\) to a state to the right of \((j,\bar{j})\). This is only possible if \(\ell \neq 0\), so that the forward probability flow due to states in stage \(i-1\) is given by

\[\sum_{\ell > 0} \pi(\ell, i-1) \lambda \sum_{m > j} \theta_{on} \]  

(52)

where the sum for \(m > j\) takes into account that only transitions with a destination on the right hand side of \((j,\bar{j})\) are to be counted.

In stage \(i\), the only states contributing some forward probability flow are \((\ell, i), \ell \leq j\) (note that all states to the right of \((j,\bar{j})\) have only upward transitions). The probability flow in this case is given by

\[\pi(0, i) \lambda \sum_{m > j} \theta_{on} + \sum_{\ell = 1}^{j} \pi(\ell, i) \lambda + (1-\lambda) \sum_{m > j} \theta_{on} \]  

(53)

where for \(\ell = 1, \cdots, j\), an upward crossing occurs with certainty if there is an arrival and only if the destination of the transition is \(m > j\) otherwise. Transitions from stage \(n\) do not cross the boundary if \(|n - \bar{n}| > 1\), so that those states do not give any contribution.

Similarly, we can find the probability flow in the reverse direction (i.e., right to left). We have

\[\sum_{\ell > j} \pi(\ell, i) (1-\lambda) \sum_{m \leq j} \theta_{on} \]  

(54)

for states in stage \(i\), and

\[\pi(0, i+1) (1-\lambda) \sum_{m \leq j} \theta_{on} \]  

(55)

for states in stage \(i+1\).

If we equate the two flows, we obtain (56), shown at the bottom of the next page, where we have defined \(\mu_{\ell j} = \sum_{m < j} \theta_{on}\).

By writing four such relationships for \(j = 0, 1, 2, 3\), we obtain the following system of equations:

\[A(\pi(1, i), \pi(2, i), \pi(3, i), \pi(0, i+1)) = B(\pi(1, i-1), \pi(2, i-1), \pi(3, i-1), \pi(0, i)) \]  

(57)

with \(A, B\) as given in (23) and (24).

If \(i = 0\), the number of possible transitions is reduced. In particular, stage \(i = 1\) does not exist, and transitions from \((0,0)\) never leave stage 0 (regardless of whether or not there is an arrival). The balance equation for state \((j,0)\) is in this case

\[\pi(0,0) (1-\mu_{0j}) = \pi(0,1) (1-\lambda) \mu_{0j} - \sum_{\ell = 1}^{j} \pi(\ell,0) (1 - (1-\lambda) \mu_{\ell j}) + \sum_{\ell > j} \pi(\ell,0) (1-\lambda) \mu_{\ell \bar{j}}. \]  

(58)

We have

\[A(\pi(1,0), \pi(2,0), \pi(3,0), \pi(0,1)) = B(0, \pi(0,0)) \]  

(59)
where $A$ is the same as before (note that the left-hand-sides of (56) and (58) are the same), whereas $B_0$ is as in (25).

**APPENDIX B**

**DERIVATION OF (31)**

The recursion (30) can be rewritten in the following way:

$$
\Phi(k, n) = \Phi(k - 1, n - 1) \begin{pmatrix} p_{00} & p_{01} \\ 0 & 0 \end{pmatrix} + \Phi(k, n - 1) \begin{pmatrix} 0 & 0 \\ p_{10} & p_{11} \end{pmatrix} + \mathbf{I}(k)\delta(n) \\
= \Phi(k - 1, n - 1)P_0 + \Phi(k, n - 1)P_1 + \mathbf{I}(k)\delta(n)
$$

where $\Phi(k, n)$ is a $2 \times 2$ matrix whose $ij$ entry is $\phi_{ij}(k, n)$.

By taking the transform

$$
\Psi(s, z) = \sum_{k=0}^{\infty} s^k \sum_{n=0}^{\infty} z^n \Phi(k, n)
$$

we obtain from (60)

$$
\Psi(s, z) = \Psi(s, z)P_0 z + \Psi(s, z)P_1 z + I
$$

which yields

$$
\Psi(s, z) = \left[ I - (P_0 s + P_1) z \right]^{-1} = \sum_{n=0}^{\infty} (P_0 s + P_1)^n z^n.
$$

The inverse transform of (63) with respect to $z$ can be found as

$$
\Phi'(s, n) = \sum_{k=0}^{\infty} s^k \Phi(k, n) = (P_0 s + P_1)^n = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}^n
$$

and the validity of (31) can be checked by direct substitution. The same result with different notation is derived in [2, p. 77].

**APPENDIX C**

**ANALYTICAL DERIVATION OF THE LATENESS PROBABILITY**

By using (9) and since $P_L(0, 0) = 0$, we can rewrite (33) as

$$
P_L = \frac{\lambda e(1-\beta)}{(1-\lambda)} \sum_{i=0}^{\infty} \beta^i P_L(0, i) + e(1-\beta) \sum_{i=0}^{\infty} \beta^i P_L(1, i)
$$

$$
= \frac{\lambda e(1-\beta)}{(1-\lambda)} \sum_{i=0}^{\infty} \beta^i P_L(0, i) + e(1-\beta) \sum_{i=0}^{\infty} \beta^i P_L(1, i)
$$

$$
= \sum_{i=0}^{\infty} \beta^i \left[ \frac{\lambda e(1-\beta)}{(1-\lambda)^2} \sum_{k=0}^{i} (\phi_{00}(k, D) + \phi_{01}(k, D)) + e(1-\beta) \sum_{k=0}^{i} (\phi_{10}(k, D) + \phi_{11}(k, D)) \right].
$$

In more compact form, by defining

$$
\pi_0 = \left( \frac{\lambda e(1-\beta)}{(1-\lambda)^2}, e(1-\beta) \right), \quad e = (1, 1)^T
$$

we have

$$
P_L(D) = \pi_0 \sum_{i=0}^{\infty} \beta^i \sum_{k=0}^{i} \Phi(k, D)e
$$

where the dependence on $D$ is now explicit in the notation. By using the recursion (60), we have (68), shown at the bottom of the page, where $M(D)$ is defined to be zero for $D < 0$. The solution of this recursion is found as

$$
M(D) = \frac{1}{1-\beta} \left[ \beta P_0 + P_1 \right]^D, \quad D \geq 0.
$$

A more explicit form can be found by observing that

$$
\pi_0[\beta P_0 + P_1] = \left( \frac{\lambda e(1-\beta)}{(1-\lambda)^2}, e(1-\beta) \right) \left( \begin{pmatrix} \beta p_{00} & \beta p_{01} \\ p_{10} & p_{11} \end{pmatrix} \right) = \frac{\epsilon(1-\beta)}{(1-\lambda)}(1-\lambda)p_{10} + \lambda p_{10}, (1-\lambda)p_{11} + \lambda p_{11} = \rho \pi_0
$$

$$
(70)
$$

$$
\sum_{i=0}^{\infty} \beta^i \sum_{k=0}^{i} \Phi(k, D) = \sum_{i=0}^{\infty} \beta^i \sum_{k=0}^{i} \Phi(k - 1, D - 1)P_0 + \sum_{i=0}^{\infty} \beta^i \sum_{k=0}^{i} \Phi(k, D - 1)P_1 + \sum_{i=0}^{\infty} \beta^i \mathbf{I}(k)\delta(D)
$$

$$
= \sum_{i=0}^{\infty} \beta^{i+1} \sum_{m=0}^{i} \Phi(m, D - 1)P_0 + \sum_{i=0}^{\infty} \beta^i \sum_{k=0}^{i} \Phi(k, D - 1)P_1 + \sum_{i=0}^{\infty} \beta^i \mathbf{I}(D)
$$

$$
= M(D - 1)[\beta P_0 + P_1] + \frac{1}{1-\beta} \mathbf{I}(D)
$$

(68)
where

\[ \rho = \frac{(1 - \lambda)p_{11} + \lambda p_{01}}{1 - \lambda} = 1 - A(\varepsilon, \lambda)p_{20} \]  

(71)

\[ A(\varepsilon, \lambda) = \frac{1 - \lambda - \varepsilon}{(1 - \lambda)(1 - \varepsilon)}. \]  

(72)

Note that the function \( A(\cdot, \cdot) \) only depends on the average failure rate \( \varepsilon \) and on the average arrival rate \( \lambda \) and is independent of the channel burstiness parameter \( p_{20} \) (in particular, it is the same as for independent and identically distributed errors).

Based on the above result, we can obtain from (69)

\[ \pi_0 M(D) = \frac{1}{1 - \beta} \pi_0 [\beta p_0 + p_z] D = \rho \pi_0 M(D - 1), \]

\[ D > 0 \]  

(73)

so that

\[ P_L(D) = \pi_0 M(D)e = P_L(1)\rho^{D-1}, \]

\[ D > 0 \]  

(74)

with

\[ P_L(1) = 1 - \pi(0, 0) = \frac{\varepsilon}{1 - \lambda}. \]  

(75)

(Note, in fact, that the only case in which the packet delay does not exceed one slot is when the incoming packet finds an empty queue and good channel.)

ACKNOWLEDGMENT

The authors would like to thank A. Cannarsi with CEFRIEL, Italy, for useful comments.

REFERENCES


Michele Zorzi (S’89–M’95–SM’98) was born in Venice, Italy, in 1966. He received the Laurea degree and the Ph.D. degree in electrical engineering from the University of Padova, Italy, in 1990 and 1994, respectively. During the academic year (1992 to 1993), he was on leave at the University of California at San Diego (UCSD) attending graduate courses and conducting research on multiple access in mobile radio networks. In 1993, he joined the faculty of the Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy. After spending three years with the Center for Wireless Communications at UCSD, in 1998, he joined the School of Engineering at the Università di Ferrara, Ferrara, Italy, where he is currently an Associate Professor. His present research interests include performance evaluation in mobile communications systems, random access in mobile radio networks, and energy constrained communications protocols.

Dr. Zorzi currently serves on the Editorial Boards of the IEEE Personal Communications Magazine and of the ACM/URSI/Baltzer Journal of Wireless Networks. He is also guest editor for special issues in the IEEE Personal Communications Magazine (Energy Management in Personal Communications Systems) and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Multimedia Network Radios).

Ramesh R. Rao (M’85–SM’90) was born in Sindri, India, in 1958. He received the bachelor’s degree (with honors) in electrical and electronics engineering from the University of Madras in 1980. He received the M.S. and Ph.D. degrees from the University of Maryland, College Park, Maryland, in 1982 and 1984, respectively.

Upon completion of the Ph.D. degree, he has been with the faculty at the Department of Electrical and Computer Engineering at the University of California at San Diego. His research interests include architectures, protocols, and performance analysis of computer and communication networks.