Performance Analysis of a Priority Queueing System over Rayleigh Fading Channels *

Min Xie and Martin Haenggi

Department of Electrical Engineering University of Notre Dame, IN 46556 USA E-mail: {mxie,mhaenggi}@nd.edu

Abstract —A wireless priority queueing system is considered for use in a delay-guaranteed sensor network, in which all sensors are transmitting towards a unique destination. In case that all sensed data in a certain period are of the same importance, a delay-balancing scheme is designed to satisfy the end-to-end delay bound. A simple priority scheduling is employed to implement this scheme. This paper statistically analyzes a two-class prioritized system with a Bounded Delay (BD) dropping strategy in a Rayleigh fading channel, modeled by a two-state Markov chain. Since the two priority classes are correlated in queueing behavior, a decomposition approach is proposed to analyze each class separately. Specifically, an approximate channel model is computed for the low priority (LP) flow so that it can be analyzed in the same way as the high priority (HP) flow(s). The QoS parameters of interest include packet delay and dropping statistics. In addition, to enable the end-to-end performance analysis, we study the approximate characterization of the output process.

1 Introduction

In the past few years, networks of low-cost, low-power, multifunctional sensors have attracted increasing attention. Different from other ad hoc wireless networks, the sensor network has the property of *light individual traffic load*, *dense deployment* and *cooperative effort of sensor nodes* [1]. Either conventional QoS-guaranteed policies for ad hoc networks have to be modified or improved, or new schemes have to be proposed to meet the QoS requirements in sensor networks.

A typical scenario occurring in sensor networks is that all sensed data are transmitted towards a unique destination (which is also referred to as sink or fusion center, see Fig. 1). Since the individual traffic load is light and could be accommodated by the wireless link, the delay tolerated by the packets at the distant nodes (the left nodes in Fig. 1) is less serious than that at the closer nodes to the sink. In other words, given the dense deployment, the nodes next to the sink (dark shaded nodes in Fig. 1) have to carry all the traffic of the network (sum of the relayed and local traffic), which may be beyond the transmission capability of the link at least temporarily, thus causing queueing delay. We refer to these nodes as *critical nodes*. Many sensor applications are delay-constrained, in particular sensor-actuator loops. A delay-balancing strategy aimed at uniformly balancing the delay is proposed so that all packets will be transmitted to

^{*}The support of the DARPA/IXO-NEST Program (AF-F30602-01-2-0526) and NSF (ECS02-25265) is gratefully acknowledged.

the sink within an identical delay bound, no matter where the packet is originated. A scheduling principle is required at least at the critical nodes to favor the relayed packets, which have experienced a longer delay than the local traffic.

Priority scheduling can be used to balance the delay of different flow packets if the relayed traffic is assigned a high priority. Moreover, channel errors due to wireless transmission (e.g., fading, path loss, shadowing) are unavoidable. The resulting priority queueing system should be analyzed using a channel model that captures the fragility of the wireless channel. This paper discusses the queueing performance of the prioritized critical nodes in a wireless Rayleigh fading channel. In addition, a Bounded Delay (BD) dropping strategy is employed, which drops the packets not sent out within the delay bound D. The QoS parameters of interest include packet delay and dropping statistics at a single node. The overall output process is also investigated for future multiple-node analysis.

The literature dealing with priority queueing systems [2–8] and wireless queueing systems [9, 10] is plentiful. A general way to characterize a pure priority queueing system is to denote the system state as the union of the queue length of each priority class and then analyze them as a whole. However, this approach does not yield the accurate probability distribution of each priority queue, particularly when the number of priorities is large. Furthermore, the channel state should be included into the system state, which results in an extra dimension to the system state space. For instance, in a single-queue system with the Rayleigh fading channel modeled by a



Figure 1: A sensor network with all nodes transmitting towards a unique sink. Darker nodes indicate higher traffic load.

simple two-state Markov chain, 2-dimensional Markov chains are required to derive the explicit steady-state probability distribution and the related dropping probability, *e.g.* [9] with {QueueLength, ChannelState}, and [10] with {HeaderPacketDelay, Channel State}. To our knowledge, no explicit results on the priority queueing system with Rayleigh fading channel have been published.

Our contribution is a decomposition approach which leads to a tractable model and approximate characterization of both the priority queues and the output process. Unlike the approaches used in wired prioritized systems, which analyze all queues jointly, our approach decouples the low priority (LP) queue from the high priority (HP) queue so that they can be analyzed separately and previous results on the single wireless queue [9,10] can be directly applied. Specifically, a virtual channel model is constructed for the LP queue which combines the impact of the HP queue and the real wireless channel. The output process is approximately characterized so that it possesses the same statistics as the input process. This approximation permits a future end-to-end (multiple node) analysis. Furthermore, this approach can be easily generalized to the multiple-queue (> 2) system in an iterative way.



Figure 2: Decomposition of the two-queue system into two single-queue systems

2 System Model

We consider a wireless link subject to bursty errors with time divided into equal-sized slots. In a slot, at most one packet can be transmitted, and each transmission may incur errors. The feedback is assumed to be error-free and available before the next packet transmission. A failed packet will be retransmitted immediately until either it is sent out successfully or its delay exceeds the delay bound D.

For bursty channel errors, a correlated error model is desired for an accurate assessment of performance. [11] and [12] have pointed out that independent channel models cannot accurately capture the system statistics. On the other hand, Wang and Chang [13] showed that a first-order Markov model is sufficient for packet level transmission in a Rayleigh fading channel. Therefore, we choose a two-state Markov chain to model the wireless channel. Let 0 and 1 denote "good" and "bad" transmission states in a given slot, respectively, and

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} \tag{1}$$

be the transition matrix for the error process, where $p_{00} + p_{01} = 1$ and $p_{10} + p_{11} = 1$. The average packet error rate is $\varepsilon = p_{01}/(p_{10} + p_{01})$.

Channel state transitions and packet arrivals occur at the end and the beginning of each slot, respectively. The arrival process is independent of the transmission errors. Both traffic flows (relayed flow indexed by 1 and the local flow 2) generate packets according to a Bernoulli process, *i.e.*, at each slot, the flow *i* has a packet arriving with a probability λ_i (i = 1, 2). Usually, we find $\lambda_1 \gg \lambda_2$ at the critical nodes. The HP packet will be transmitted immediately if its queue is empty, while the LP packet has to wait until the HP queue is empty. In each queue, the packets are served in a FIFO manner.

With these assumptions, the HP queue is exclusively served by the channel and independent of the LP queue, while the LP queue is served only when the HP queue does not occupy the channel. The wired and continuous counterpart on this queueing system is given in [14] using fluid flow models. Following a similar principle, we decouple our discrete wireless priority queueing system.

3 Decomposition Approach

In Fig. 2, the arrival process and the output process are denoted by $a_i(t)$ and $d_i(t)$, respectively. The channel rate c(t) is modulated by a two-state Markov chain (1). The system output process d(t) is the sum of the two individual output processes, $d(t) = d_1(t) + d_2(t)$. As discussed in Section 2, the HP queue can be analyzed as a single-queue system with a channel rate c(t), and the LP queue is served by the remaining service $v(t) = c(t) - d_1(t)$. Note that if we let the state space of the output process be $\{0, 1\}$ (to be consistent with the state space of the channel) in which 1 represent a packet departing



Figure 3: Sample path of the two-priority system

the system, and 0 otherwise, then the service rate of the LP queue is

$$v(t) = (c(t) - d_1(t)) \mod 2 = c(t) + d_1(t).$$
 note that $c(t)d_1(t) \equiv 0.$ (2)

The mod 2 operation ensures that the state space of v(t) is $\{0, 1\}$. Fig. 3 shows a sample path of each process. Since v(t) has the same state space as the original wireless channel c(t), and the LP packet is transmitted successfully only when v(t) = 0 (see Fig. 3), we call v(t) the rate of a *virtual channel* for the LP queue. These decoupled single-queue systems (in Fig. 2) with a BD policy can be analyzed using the approach introduced in [10].

Our objective is to characterize the virtual channel. As shown in (2) and Fig. 3, v(t) = 0 if c(t) = 0 and $d_1(t) = 0$, *i.e.*, when the channel is good and no HP packet departs. $d_1(t)$ is unknown, but from $c(t)d_1(t) \equiv 0$ we conclude that $d_1(t)$ is correlated to c(t) and can be derived from the HP system.

3.1 A Single Wireless Queue with a BD Strategy

[10] has developed a model to describe the BD dropping system, in which a twodimensional Markov chain with state space $\{(i, j) | -\infty < i \leq D, j = 0, 1\}$ is used, where D is the maximum tolerable packet delay. The first entry i keeps track of the packet delay, and the second models the error process. When i < 0 (negative delay), the queue is empty and |i| represents the remaining time before a packet arrives at the empty queue.

In the HP system, let X(n) denote the system state at time n and D_1 the HP delay bound. Define $P_{(i,j),(k,l)}$ as the one-step system transition probability

$$P_{(i,j),(k,l)} = \Pr\{X(n+1) = (k,l) \mid X(n) = (i,j)\} \\ = \begin{cases} p_{jl} & i < 0, \ k = i+1 \\ a_1^{i-k+1}p_{0l} & 0 \le i \le D_1, \ k \le i, \ j = 0 \\ p_{1l} & 0 \le i < D_1, \ k = i+1, \ j = 1 \\ a_1^{D_1-k+1}p_{1l} & i = D_1, \ k \le D_1, \ j = 1 \\ 0 & \text{otherwise}, \end{cases}$$
(3)

where

$$a_1^t := \lambda_1 (1 - \lambda_1)^{t-1}$$
; (note that $\sum_{t=1}^{\infty} a_1^t = 1$). (4)

The resulting Markov chain is of infinite length. The steady-state probabilities $\pi(i, j)$ are solved by using the balance equations of the Markov chain,

$$\pi(i,j) = \begin{cases} \frac{p_{10}}{\omega} x^{D_1 - i} \pi(D_1, 1) & i > 0, \ j = 0\\ (1 - \lambda_1)^{D_1 - i} \Big[1 + \frac{p_{10}(1 - \omega)^{D_1 - i}}{(1 - \omega)\omega^{D_1 - i}} \Big] \pi(D_1, 1) & i \ge 0, \ j = 1\\ \frac{p_{10}}{p_{01}} x^{D_1} \pi(D_1, 1) & i = 0, \ j = 0\\ (1 - \lambda_1)^{-i} \frac{p_{10}}{p_{01}} x^{D_1} \pi(D_1, 1) & i < 0, \ j = 0\\ (1 - \lambda_1)^{D_1 - i} (1 + p_{10}\omega_{D_1}) \pi(D_1, 1) & i < 0, \ j = 1, \end{cases}$$

$$(5)$$

where

$$\begin{cases} \omega = 1 - \lambda_1 (1 - p_{01}) - p_{10} (1 - \lambda_1) \\ x = \frac{1 - \lambda_1}{\omega} \\ \omega_{D_1} = \frac{1 - \omega^{D_1}}{\omega^{D_1} (1 - \omega)}. \end{cases}$$

Note that the first three expressions in (5) have been derived in [10], and the last two equations are obtained after we apply the condition of *infinite interarrival time* $(A_1 = \infty)$ to $\{\pi(i, j) \mid i < 0\}$.

The probability $\pi(D_1, 1)$ is calculated by using

$$\sum_{i=-\infty}^{D_1} \sum_{j=0}^{1} \pi(i,j) = 1 \quad ; \quad \text{therefore} \quad \pi(D_1,1) = \frac{1}{k_{sum}}, \tag{6}$$

where

$$k_{sum} = \frac{p_{10}x^{D_1}}{p_{01}\lambda_1} + (1+p_{10}\omega_{D_1})\frac{(1-\lambda_1)^{D_1+1}}{\lambda_1} + \frac{\left[1-(1-\lambda_1)^{D_1+1}\right](1-\omega-p_{10})}{\lambda_1(1-\omega)} + \frac{p_{10}(1-x^{D_1})}{\omega(1-x)} + \frac{p_{10}(1-x^{D_1+1})}{(1-\omega)(1-x)}.$$
(7)

3.2 Characterization of the Virtual Channel v(t)

We find that as i < 0 ($\Leftrightarrow d_1(t) = 0$) and j = 0 ($\Leftrightarrow c(t) = 0$), there is no packet in the HP queue (thus no packet departure) and the channel is good, which corresponds to v(t) = 0; the remaining states correspond to v(t) = 1. Since the virtual channel cares only about the transition between 0 and 1, not how the HP system evolves, it is sensible to divide the HP state space into two subsets J and B, defined as $J = \{(i, j) \mid i < 0 \text{ and} j = 0\}$, $\mathcal{B} = \mathcal{I}^C$. The subset J and \mathcal{B} correspond to 0 and 1 of v(t), respectively. Now, we compute the transition probabilities between these two subsets. (Note that we use \sim , as in \tilde{p}_{10} and $\tilde{\varepsilon}$, to denote the parameters associated with the virtual channel v(t)).

$$\tilde{p}_{01} = \Pr\{X(n+1) \in \mathcal{B} | X(n) \in \mathcal{I}\} = \frac{\Pr\{X(n+1) \in \mathcal{B}, X(n) \in \mathcal{I}\}}{\Pr\{X(n) \in \mathcal{I}\}}$$

$$= \frac{\sum_{i < 0} \pi(i, 0) \left(\sum_{k = -\infty}^{D_1} P_{(i,0),(k,1)} + \sum_{k = 0}^{D_1} P_{(i,0),(k,0)}\right)}{\sum_{i < 0} \pi(i, 0)}$$

$$= p_{01} + \frac{\pi(-1, 0)}{\sum_{i < 0} \pi(i, 0)} (1 - p_{01}) = p_{01} + \lambda_1 (1 - p_{01}); \quad (8)$$

$$\tilde{p}_{10} = \Pr\{X(n+1) \in \mathcal{I} | X(n) \in \mathcal{B}\} = \frac{\Pr\{X(n+1) \in \mathcal{I}, X(n) \in \mathcal{B}\}}{\Pr\{X(n) \in \mathcal{B}\}}$$
$$= \frac{\sum_{i=-\infty}^{D_1} \sum_{k<0} \pi(i,1) P_{(i,1),(k,0)} + \sum_{i=0}^{D_1} \sum_{k<0} \pi(i,0) P_{(i,0),(k,0)}}{\sum_{i=-\infty}^{D_1} \pi(i,1) + \sum_{i=0}^{D_1} \pi(i,0)} = \frac{k_{10}}{k_1},$$
(9)

where

$$k_{10} = p_{10}(1-\lambda_1)^{D_1+1} \Big[(1-p_{01}-p_{10})\omega_{D_1} + \frac{1+p_{10}\omega_{D_1}}{\lambda_1} + \frac{1-p_{01}}{p_{01}\omega^{D_1}} \Big] \pi(D_1,1)$$

$$k_1 = 1 - \frac{p_{10}}{p_{01}} \frac{1-\lambda_1}{\lambda_1} x^{D_1} \pi(D_1,1).$$

 \tilde{p}_{01} in (8) is consistent with our intuition. There is a transition from v = 0 to v = 1 if either the channel becomes bad (*i.e.*, $j : 0 \to 1$ with probability p_{01}) or a packet arrives but the channel remains good (*i.e.*, $i : -1 \to 0, j = 0$ with probability $\lambda_1 p_{00}$).

Furthermore, the average error rate $\tilde{\varepsilon}$ is obtained in a straightforward way:

$$\tilde{\varepsilon} = \frac{\tilde{p}_{01}}{\tilde{p}_{01} + \tilde{p}_{10}} = 1 - \frac{p_{10}}{p_{01}} \frac{1 - \lambda_1}{\lambda_1} x^{D_1} \pi(D_1, 1) = \sum_{i=0}^{D_1} \pi(i, 0) + \sum_{i=-\infty}^{D_1} \pi(i, 1) = \Pr\{X \in \mathcal{B}\}.$$
(10)

The computed error rate $\tilde{\varepsilon}$ is identical to the sum probability of all bad states in \mathcal{B} , which proves that our approach of state aggregation is reasonable. In summary, we aggregate all the good states of the HP queueing system into one state 0 of the virtual channel v(t) and all the bad states into 1. Thus we have another two-state Markov model $\{\tilde{p}_{ij}|i, j = 0, 1\}$ for v(t).

3.3 The Two-Priority Queueing System

For a two-priority queueing system, using the above decomposition approach, we obtain two single-queue systems, whose channel processes are c(t) and v(t), modulated by twostate Markov chains with transition probabilities $\{p_{ij} \mid i, j = 0, 1\}$ and $\{\tilde{p}_{ij} \mid i, j = 0, 1\}$, respectively. Each class has its own delay bound D_k (k = 1, 2). A reasonable assumption is that $D_2 > D_1$ because the locally generated packets can tolerate a longer delay than the relayed packets if they have the same end-to-end delay constraint. Naturally, the packet delay $d_j^{(k)}$ and packet loss probability $p_L^{(k)}$ for queue k are cal-

Naturally, the packet delay $d_j^{(\kappa)}$ and packet loss probability $p_L^{(\kappa)}$ for queue k are calculated as in the single-queue system (given in [10]):

$$d_{j}^{(k)} = \frac{\pi^{(k)}(j,0)}{\sum_{i=0}^{D_{k}} \pi^{(k)}(i,0)} \quad \text{and} \quad p_{L}^{(k)} = \frac{\pi^{(k)}(D_{k},1)}{\lambda_{k}}, \quad (11)$$

where $\pi^{(1)}(i,j) = \pi(i,j)$ in (5), and $\pi^{(2)}(i,j)$ is calculated in the same way as in (5), but with different parameters λ_2 and $\{\tilde{p}_{ij}|i,j=0,1\}$ given by (8) and (9).

With this decomposed system, it is not difficult to approximate the output process d(t). The arrival processes are assumed to be Bernoulli (the interarrival time is geometrically distributed and memoryless). However, due to the correlations in the channel errors and the two queues, the output process does not preserve this memoryless property. In order to carry out the end-to-end analysis, it is desired that the output process has the same probabilistic properties as the input process. A natural idea is to introduce some approximations and make the output process memoryless. The key in the Bernoulli process is the probability λ of an event occurring. In terms of the output process, the wanted probability is the average departure probability.

As pointed out before, the output process d(t) is the sum of two individual processes $d_1(t)$ and $d_2(t)$. Denote the departure probability of queue k by $p_d^{(k)}$. By means of the decomposition principle, the overall departure probability is:

$$p_d = p_d^{(1)} + p_d^{(2)} (1 - \tilde{\varepsilon}).$$
(12)

Similar to the calculation of $\tilde{\varepsilon}$, the $p_d^{(k)}$ can be calculated by dividing the system space into two subsets. For instance, in the HP system, a packet departs when $i \ge 0$ and j = 0. Consequently, we denote the subset $\mathcal{D} = \{(i, j) \mid i \ge 0 \text{ and } j = 0\}$ as the set of all states in which a packet departure occurs, and $\mathcal{N} = \mathcal{D}^C$ is the set of remaining states. Therefore, the departure probability for the HP queue is given by the equation

$$p_d^{(1)} = \Pr\{X \in \mathcal{D}\} = \sum_{i=0}^{D_1} \pi(i,0) = \frac{p_{10}}{p_{01}} x^{D_1} \left(1 + \frac{p_{01}}{\omega} \frac{1 - x^{D_1}}{1 - x}\right) \pi(D_1,1).$$
(13)

 $p_d^{(2)}$ is computed in the same way with the parameters and probabilities of the LP queue.

4 Numerical Results

In this section, we present some example results obtained based on the analysis described above. Since the focus is on the critical nodes, the exported traffic load λ_1 is assumed to be heavy, and $\lambda_2 \ll \lambda_1$. As mentioned in Section 3.3, $D_2 > D_1$. Specifically, we set $D_2 = 10D_1$. We first display the impact of the delay bound on the packet loss probability. The arrival rates are $\lambda_1 = 0.8$ and $\lambda_2 = 0.08$, respectively; the mean burst error length is $\frac{1}{p_{10}} = 10$, and the average error rate $\varepsilon = 10^{-2}$. This setting implies a heavy traffic load and highly correlated errors, which is the worst case system. Fig. 4 exhibits the dropping probabilities p_L of the two queues. Besides the part where the delay bound is very small ($D_1 \leq 4$ which is not practical — in fact, we usually have $D_1 \geq 10$), the



Figure 4: Packet loss probability vs. delay bound D_1 with $D_2 = 10D_1$



(a) Packet loss probability vs. channel transition probability p_{10}



Figure 5: Packet loss probability vs. channel error rate ε



(b) Error rate of the virtual channel vs. channel transition probability p_{10}

Figure 6: The impact of the channel transition probability p_{10}

packet loss probability of the local traffic is much smaller than that of the relayed traffic (can be ignored when $D_1 \ge 20$). This is consistent with our expectation. Even though the relayed traffic exclusively occupies the channel, most local packets are still sent out successfully, with a fairly small loss. Therefore, the delay-balancing scheme improves the delay of the long-path packets at a small price on the short-path packets. Moreover, as proved by [10], the packet loss probability (log scale) is linear to the delay bound.

Next, we investigate the impact of the channel. First consider the channel error rate ε . The delay bounds are fixed at $D_1 = 10$ and $D_2 = 100$, and other parameters remain unchanged. Fig. 5 displays the two queues' packet loss probability versus the channel error rate ε . The HP queue is like a single queue with a server c(t), and its $p_L^{(1)}$ (log scale) is near-linear to $\log(\varepsilon)$ (which is identical to the result of [10]). On the other hand, for LP packets, when ε is small (< 10^{-3}), the packet loss probability is almost independent of ε . Even as ε increases (to 10^{-2}), the change of $p_L^{(2)}$ is still very small.

 p_{10} is another channel parameter considered. Fig. 6(a) shows that, like ε , p_{10} does not play a significant role in $p_L^{(2)}$. Fig. 6(b) shows how the error rate $\tilde{\varepsilon}$ of the virtual channel changes with p_{10} . No apparent change in $\tilde{\varepsilon}$ happens when the channel error correlation decreases (p_{10} increases). That explains why $p_L^{(2)}$ does not greatly change with p_{10} .



(a) Packet loss probability vs. arrival rate λ with $\lambda_1+\lambda_2=0.9$



(b) Error rate of the virtual channel $\tilde{\varepsilon}$ vs. arrival rate λ

Figure 7: The impact of the HP flow's arrival rate λ_1

The last one is the arrival rate λ_i . Fig. 7(a) and Fig. 7(b) show the influence of λ_i given that $\lambda_1 + \lambda_2 = 0.9$. The virtual channel error rate $\tilde{\varepsilon}$ linearly increases with λ_1 , which results in the increase of the LP queue's packet loss probability $p_L^{(2)}$ (Fig. 7(a)).

In summary, the LP queue's performance is determined by the virtual channel, which, in turn, is a combination of the HP queue and the real channel. Our numerical results show that the HP flow is more important for the LP queue than the channel c(t). This is reasonable, since the channel is good on average (given the assumption of small error rate ε), and the bad period of the virtual channel is highly possibly caused by the arrival of HP packets. Therefore, small changes in the arrival rate of the HP flow will lead to a bigger improvement of the LP packet loss probability, compared to the channel statistics.

5 Conclusions and Future Work

A decomposition approach is proposed to solve the queueing problems in a priority queueing system with correlated channel errors and real-time traffic. Closed-form solutions for each queue's delay distribution and dropping probability are provided by partitioning the HP system state space into two subsets and aggregating all the states in a subset to achieve a two-state Markov model for the LP queue. The numerical results give some insights on the design of the system. When the wireless channel is not bad on average (small error rate), which is reasonable in practice, the queueing performance of low priorities is mainly determined by that of the high priorities, rather than the channel itself. Moreover, if the delay bounds are chosen appropriately (*e.g.* $D_2 = 10D_1$) and the LP traffic load is light, the loss rate in the LP packets is small.

This approach can be easily generalized to a k-priority system (k > 2) in an iterative way by constructing a virtual channel for each lower priority class i (i > 1). Our future work includes the extensions from two aspects. First, extend from the single-node case to the multiple-node case; Second, extend from the simple Bernoulli arrival process to more general arrival processes.

References

- I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Networks," *IEEE Communications Magazine*, vol. 40, pp. 102–114, Aug. 2002.
- [2] G. Barberis, "A Useful Tool in the Theory of Priority Queueing," *IEEE Transactions on Communications*, vol. 28, pp. 1757–1762, Sept. 1980.
- [3] J. N. Daigle and M. Roughan, "Queue-Length Distributions for Multi-Priority Queueing Systems," in *IEEE INFOCOM*, vol. 2, pp. 21–25, Mar. 1999.
- [4] D. A. Gall and H. R. Mueller, "Waiting-Time Distributions and Buffer Overflow in Priority Queueing Systems," *IEEE Transactions on Communications*, vol. 20, pp. 865–877, Oct. 1972.
- [5] D. McMillan, "Delay Analysis of a Cellular Mobile Priority Queueing System," IEEE/ACM Transactions on Networking, vol. 3, pp. 310–319, June 1995.
- [6] A. Khamisy and M. Sidi, "Discrete-Time Priority Queues with Two-State Markov Modulated Arrivals," *Stochastic Models*, vol. 8, no. 2, pp. 337–357, 1992.
- [7] E. Gelenbe and A. Ghanwani, "Approximate Analysis of Coupled Queueing in ATM Networks," *IEEE Communications Letters*, vol. 3, pp. 31–33, Feb. 1999.
- [8] R. Fantacci, "Performance Evaluation of Prioritized Handoff Schemes in Mobile Cellular Networks," *IEEE Transactions on Vehicular Technology*, vol. 49, pp. 485–493, Mar. 2000.
- [9] M. Zorzi, "Data-Link Packet Dropping Models for Wireless Local Communications," IEEE Transactions on Vehicular Technology, vol. 51, pp. 710–719, July 2002.
- [10] K. K. Lee and S. T. Chanson, "Packet Loss Probability for Real-Time Wireless Communications," *IEEE Transactions on Vehicular Technology*, vol. 51, pp. 1569–1575, Nov. 2002.
- [11] D.-L. Lu and J.-F. Chang, "Performance of ARQ Protocols in Nonindependent Channel Errors," *IEEE Transactions on Communications*, vol. 41, pp. 721–730, May 1993.
- [12] M. Zorzi, "Packet Dropping Statistics of a Data-Link Protocol for Wireless Local Communications," in *IEEE International Conference on Universal Personal Communications* (*ICUPC*), pp. 536–540, Oct. 1997.
- [13] H. S. Wang and P.-C. Chang, "On Verifying the First-Order Markovian Assumption for a Rayleigh Fading Channel Model," *IEEE Transactions on Vehicular Technology*, vol. 45, pp. 353–357, May 1996.
- [14] A. Elwalid and D. Mitra, "Analysis, Approximations and Admission Control of a Multi-Service Multiplexing System with Priorities," in *IEEE INFOCOM95*, pp. 463–472, 1995.