A Simple and Effective Priority Scheme for IEEE 802.11

Yang Xiao, *Member, IEEE*

***Abstract*—In this letter, we propose an analytical model for a simple priority scheme for real-time applications in IEEE 802.11 by differentiating the initial window size, the window-increasing factor and the maximum backoff stage. Saturation throughputs and saturation delays of different priority classes are derived analytically.**

***Index Terms*—IEEE 802.11, performance analysis, priority.**

I. Introduction

The IEEE 802.11 distributed coordination function (DCF) is a very robust protocol for the best-effort service in the wireless medium. However, it is unsuitable for real-time applications. One possible solution is to provide a good priority scheme for DCF. Deng and Chang [2] proposed a priority scheme by differentiating the backoff window: the higher priority class uses the window $[0, 2^{j+1}-1]$ and the lower priority class uses the window $[2^{j+1}, 2^{j+2}-1]$, where j is the backoff stage. Aad and Castelluccia [3] proposed a priority scheme achieved by differentiating inter-frame spaces (IFSs). Veres and Campbell *et al.* [4] proposed priority schemes by differentiating the initial backoff window size and the window size. Pallot and Miller [8] proposed three priority schemes: static priority scheduling, prioritized DIFS time mechanism and prioritized backoff time distribution mechanism (PBTDM). PBTDM is a very interesting approach in which the backoff time is chosen in the current window range with different distributions for different priorities. All the priority schemes [2]–[4], [8] are conducted based on simulations.

There have been many analytical models proposed for IEEE 802.11 performance analysis. Bianchi [5], [9] proposed an accurate analytical model to compute saturation throughput and Ziouva *et al.* [6] improved Bianchi's model by deriving saturation delay. None of their models are for priorities.

In this letter, based on Bianchi's [9] and Ziouva's [6] models, we propose an analytical model for a simple priority scheme. The advantage of our model is to provide priorities.

II. The Priority Scheme

Assume that traffic is classified into N priority classes: $i = 1, \ldots N$. We modify 802.11 MAC to provide a priority scheme by differentiating following three metrics for the priority i class: the initial window size $W_{i,0}$, the window-increasing factor σ_i and the maximum backoff stage m_i, where σ_i is the factor by which the current window size is increased when a transmitted frame collides and equals 2 in IEEE 802.11 [1]. In this letter, we let σ_i be a real number and $\sigma_i > 1$. If assuming that the priority i class has higher priority than the priority j class, we have $W_{i,0} \le W_{j,0}$, $1 < \sigma_i \le \sigma_j$ and $m_i \le m_j$. Furthermore, at least one of the above inequalities must be a real inequality. If one class has a smaller metric, the class's traffic has a better chance to access the channel earlier. Many special cases of our scheme can be designed by differentiating any one or two metrics among the three metrics.

Manuscript received Auguest 24, 2002. The associate editor coordinating the review of this letter and approving it for publication was Prof. C. Douligeris.

The author is with Computer Science Division, The University of Memphis, Memphis, TN 38152 USA (e-mail: yangxiao@ieee.org).

Digital Object Identifier 10.1109/LCOMM.2002.808370

III. Performance Evaluation

A. An Analytical Model

Assume that each station belongs to only one priority class and always has frames ready to send. For a given station in the priority i class, $b(i,t)$ is defined as a random process representing the value of backoff counter at time t and $s(i,t)$ is defined as the random process representing the backoff stage j, where $0 \le j \le m_i$. The value of $b(i,t)$ is uniformly chosen in the range $(0, 1, \ldots W_{i,j}-1)$, where $W_{i,j} = \lceil (\sigma_i)^j W_{i,0} \rceil$. Let p_i denote the probability that a transmitted packet collides and p_b denote the probability that the channel is busy. Similar to Bianchi's model [5], [9] and Ziouva's model [6], the bi-dimensional random process $\{s(i,t), b(i,t)\}$ is discrete-time Markov chain under the assumptions that p_i and p_b are both independent to the backoff procedure. Therefore, the state of each station in the priority i class is described by $\{i,j,k\}$, where j stands for the backoff stage taking values from $(0, 1, \ldots m_i)$ and k stands for the backoff delay taking values from $(0, 1, \ldots W_{i,j}-1)$ in timeslots. The state transition diagram for the priority i class is shown in Fig. 1 with the nonnull transition probabilities, where the state $\{i, -1, 0\}$ stands for the state that the station senses the channel idle after DIFS and transmits successfully without activating the backoff stage [6].

Let $b_{i,j,k} = \lim_{t\to\infty} \Pr\{s(i,t) = j, b(i,t) = k\}$ be the stationary distribution of the Markov chain. In steady state, we can derive following relations through chain regularities:

$$b_{i,j,0} = p_i^j b_{i,0,0}, \qquad 0 \le j \le m_i - 1 \tag{1}$$

$$b_{i,m_i,0} = \frac{p_i^{m_i} b_{i,0,0}}{(1-p_i)} \tag{2}$$

$$b_{i,j,k} = \frac{W_{i,j}-k}{W_{i,j}} \frac{1}{1-p_b} b_{i,j,0}, \qquad 0 \le j \le m_i, 1 \le k \le W_{i,j}-1 \tag{3}$$

$$b_{i,0,0} = \frac{1-(1-p_i)(1-p_b)}{1-p_b} b_{i,-1,0} \tag{4}$$

$$b_{i,-1,0} + \sum_{j=0}^{m_i} \sum_{k=0}^{W_{i,j}-1} b_{i,j,k} = 1. \tag{5}$$

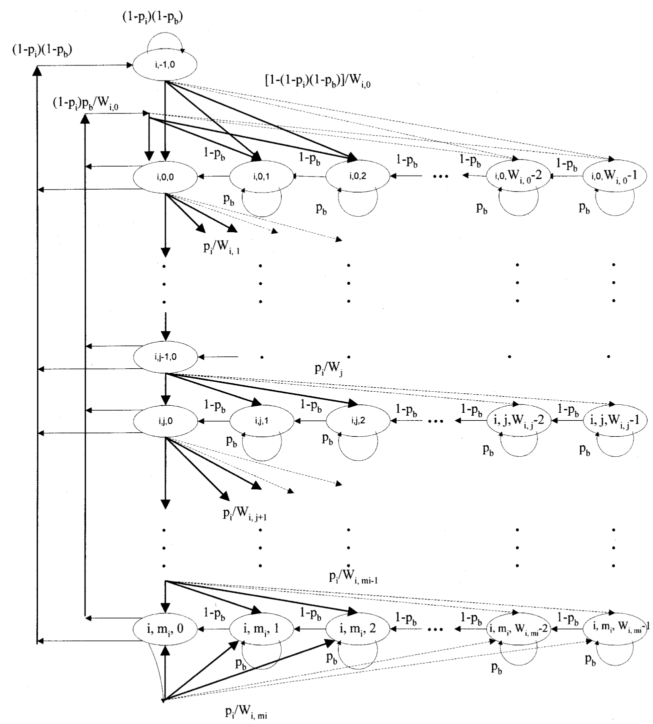

Fig. 1. The state transition diagram for the priority i class.

Let $\alpha_1 = 2(1-p_b)^2(1-p_i)(1-p_i\sigma_i)$, $\alpha_2 = [1-(1-p_i)(1-p_b)]$, and $\alpha_3 = [(1-\sigma_i p_i) + (1 - p_i + \sigma_i^{m_i} p_i^{m_i+1} - p_i^{m_i+1}\sigma_i^{m_i+1})W_{i,0}]$. Let τ_i be the probability that a station in the priority i class transmits during a generic slot time. Let $n_i (i = 1, \ldots, N)$ denote the number of stations in the priority i class. We have

$$b_{i,-1,0} = \frac{\alpha_1}{\alpha_1 + \alpha_2\alpha_3} \quad (6)$$

$$\tau_i = \sum_{j=-1}^{m_i} b_{i,j,0} = \frac{2(1-p_b)(1-\sigma_i p_i)}{\alpha_1 + \alpha_2\alpha_3} \quad (7)$$

$$p_b = 1 - \prod_{h=1}^{N}(1-\tau_h)^{n_h} \quad (8)$$

$$p_i = 1 - \left(\prod_{h=1}^{i-1}(1-\tau_h)^{n_h}\right)(1-\tau_i)^{n_i-1} \times \left(\prod_{h=i+1}^{N}(1-\tau_h)^{n_h}\right). \quad (9)$$

Substituting (8) and (9) in (7), we can solve them numerically. Then, we can calculate p_b and p_i from (8) and (9).

B. Saturation Throughput

Let $p_{s,i}(i = 1, \ldots N)$ denote the probability that a successful transmission occurs in a slot time for the priority i class and let p_s denote the probability that a successful transmission occurs in a slot time. Therefore, p_s/p_b is the probability that the transmitted frame is successful. We have

$$p_{s,i} = n_i\tau_i(1-\tau_i)^{n_i-1}\prod_{h=1,h\neq i}^{N}(1-\tau_h)^{n_h} = \frac{n_i\tau_i}{1-\tau_i}(1-p_b) \quad (10)$$

$$p_s = \sum_{i=1}^{N} p_{s,i} = (1-p_b)\sum_{h=1}^{N}\frac{n_h\tau_h}{1-\tau_h}. \quad (11)$$

Let $S_i (i = 1, \ldots N)$ denote the normalized throughput for the priority i class. Let δ, $T_{E(L)}$, T_s and T_c denote the duration of an empty slot time, the time to transmit the payload, the average time that the channel is sensed busy because of a successful transmission and the average time that the channel has a collision, respectively. We have

$$S_i = \frac{E(\text{payload transmission time in a slot time for the } i \text{ class})}{E(\text{length of a slot time})} = \frac{p_{s,i}T_{E(L)}}{(1-p_b)\delta + p_sT_s + [p_b - p_s]T_c} \quad (12)$$

If $N = 1$ (only one class), it is easy to prove that (12) is equivalent to (13) in [9], or (9) in [6], although notations are a little different.

Let $T_H, T_{ACK}, SIFS, DIFS, L^*, T_{E(L^*)}$ and γ denote the time to transmit the header, the time to transmit the ACK, SIFS time, DIFS time, the length of the longest packet in a collision, the time to transmit a payload with length $E(L^*)$ and the propagation delay, respectively. We have

$$T_s = T_H + T_{E(L)} + SIFS + \gamma + T_{ACK} + DIFS + \gamma \quad (13)$$

$$T_c = T_H + T_{E(L*)} + DIFS + \gamma. \quad (14)$$

C. Saturation Delay

For the priority i class, let $N_{c,i}$ denote the random variable representing the number of collisions before transmitting a frame; let X_i denote the random variable representing the time interval during which the counter reaches zero without considering the case when the counter freezes; let F_i denote the time that the backoff counter of a station freezes; let N_{F_i} denote the number of times that the backoff counter freezes; let B_i denote the backoff delay of a station before accessing the channel under busy channel condition; let D_i denote the random variable representing the frame delay; let T_o denote the time that a station has to wait when its frame transmission collides before sensing the channel again; let T_{timeout} denote the duration of the ACK timeout. We have

$$E(N_{c,i}) = \frac{p_b}{p_{s,i}} - 1 \quad (15)$$

$$E(X_i) = \sum_{j=0}^{m_i}\sum_{k=1}^{W_{i,j}-1} k b_{i,j,k} \quad (16)$$

$$E(N_{F_i}) = \frac{E(X_i)}{\max\left(\frac{1-p_b}{p_b}, 1\right)} - 1 \quad (17)$$

$$E(F_i) = E(N_{F_i})\left(\frac{p_s}{p_b}T_s + \left(1 - \frac{p_s}{p_b}\right)T_c\right) \quad (18)$$

$$E(B_i) = E(X_i) + E(F_i) \quad (19)$$

$$E(D_i) = E(N_{c,i})[E(B_i) + T_c + T_o] + E(B_i) + T_s \quad (20)$$

$$T_o = SIFS + T_{\text{timeout}} \quad (21)$$

D. Numerical Results

We use IEEE 802.11a as an example. Both the data rate and the control rate are 6 Mb/s. The packet size is 1024 bytes. The

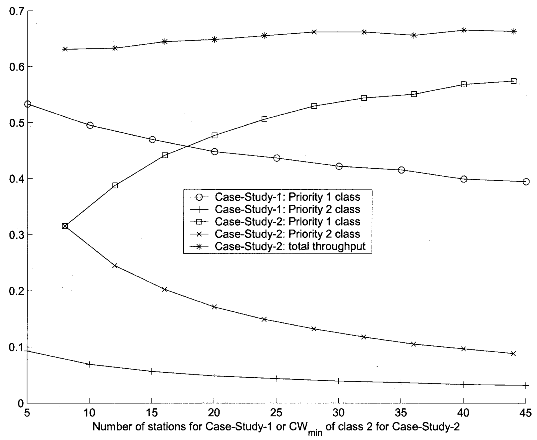

Fig. 2. Saturation throughput (normalized).

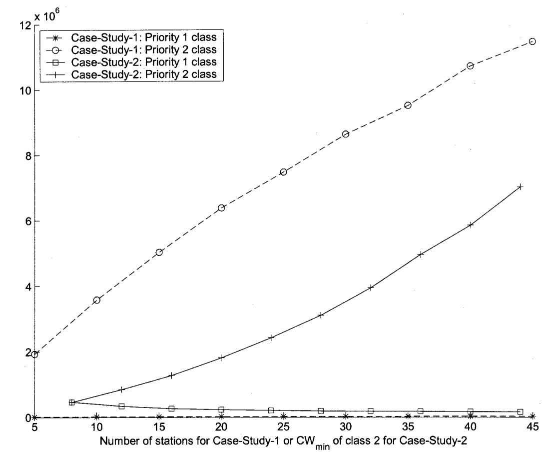

Fig. 3. Saturation delay (μ seconds.

parameters for IEEE 802.11a can be found in [7], as well as how to calculate $T_H+T_{E(L)}$ accurately. For demonstration purposes, we adopt two priority classes, i.e., $N = 2$. However, our proposed model is very general so that we can design many levels of priorities. Figs. 2 and 3 show the saturation throughput and saturation delay, respectively, over number of stations (n_1 or n_2) for the case study one and over $W_{2,0}$ for the case study two. Note that for different case studies, x-axis stands for different things to save space.

For the case study one, following parameters are adopted: $[\sigma_1, \sigma_2] = [1.7, 2]$, $[m_1, m_2] = [4, 7]$, $[W_{1,0}, W_{2,0}] = [4, 8]$ and $n_1 = n_2$. The class 1 has a much better throughput (delay) than the class 2. Fig. 3 also indicates that the delay for class 1 is very small (in magnitude of 10^4). One application of such a simple example is to use the class 1 for real-time applications and to use the class 2 for best-effort applications. These figures indicate that the proposed priority scheme is very effective.

For the case study 2, following parameters are adopted: $\sigma_1 = \sigma_2 = 2$, $m_1 = m_2 = 7$, $W_{1,0} = 8$ and $n_1 = n_2 = 10$. When $W_{2,0} = 8$, throughputs (delays) are the same for both classes. As $W_{2,0}$ increases, the throughput of class 2 decreases, the throughput of class 1 increases, the delay of class 1 decreases and the delay of class 2 increases dramatically. Therefore, the delay of class 2 is very sensitive to $W_{2,0}$. An interesting observation in Fig. 2 is that the throughputs of class 1 and class 2 are symmetric along a line parallel to the x-axis. This phenomenon indicates that class 1 can steal throughput from class 2 as $W_{2,0}$ increases, whereas the total throughput of all classes does not change much due to the following reasons. As $W_{2,0}$ increases, stations in class 2 will delay accessing the channel so that the throughput of class 2 will decrease and the delay of class 2 will increase. Furthermore, collision probabilities of class 1 will decrease so that the throughput of class 1 will increase and the delay of class 1 will decrease. We also observe that in Fig. 3, as $W_{2,0}$ increases, the delay of class 1 decreases only a little.

Similar numerical results (omitted due to limited space) show that any of the proposed three metrics can provide good priorities among classes and provide good service differentiations. For all three metrics, one class can steal bandwidth from another if the later one increases the metric value and the total throughput does not change much. This fact indicates that they are good metrics. However, their delays cannot get a lot of benefits. Three metrics can be implemented all together since this way does not make a hardware implementation more difficult.

ACKNOWLEDGMENT

The author would like to thank Dr. C. Douligeris and the three anonymous reviewers for their valuable comments that have greatly improved the quality of this letter.

REFERENCES

[1] *IEEE 802.11 WG Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, Standard, IEEE*, 1999.
[2] D.-J. Deng and R.-S. Chang, "A priority scheme for IEEE 802.11 DCF access method," *IEICE Trans. Commun.*, vol. E82-B, no. 1, Jan. 1999.
[3] I. Aad and C. Castelluccia, "Differentiation mechanisms for IEEE 802.11," in *IEEE INFOCOM 2001*, pp. 209–218.
[4] A. Veres, A. T. Campbell, M. Barry, and L.-H. Sun, "Supporting differentiation in wireless packet networks using distributed control," *IEEE J. Select, Areas Commun.*, vol. 19, pp. 2081–2093, Oct. 2001.
[5] G. Bianchi, "IEEE 802.11—Saturation throughput analysis," *IEEE Commun. Lett.*, vol. 2, pp. 318–320, Dec. 1998.
[6] E. Ziouva and T. Antonakopoulos, "CSMA/CA performance under high traffic conditions: Throughput and delay analysis," *Comp. Commun.*, vol. 25, pp. 313–321, 2002.
[7] Y. Xiao and J. Rosdahl, "Throughput and delay limits of IEEE 802.11," *IEEE Commun. Lett.*, vol. 6, pp. 355–357, Aug. 2002.
[8] X. Pallot and L. E. Miller, "Implementing message priority policies over an 802.11 based mobile ad hoc network," *IEEE MILCOM 2001*, pp. 861–865.
[9] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 318–320, Mar. 2000.