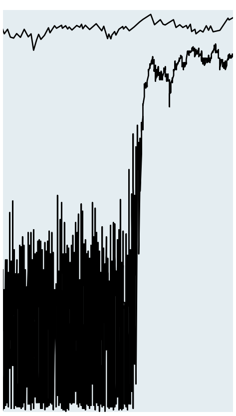
PROVIDING STATISTICAL QOS GUARANTEE FOR VOICE OVER IP IN THE IEEE 802.11 WIRELESS LANS

HONGQIANG ZHAI, JIANFENG WANG, AND YUGUANG FANG, UNIVERSITY OF FLORIDA



The authors propose a call admission and rate control scheme to provide statistical QoS guarantees for VoIP calls. Extensive simulations demonstrate that the proposed schemes can support statistical QoS guarantees for voice traffic and maintain high throughput for non-voice traffic at the same time.

ABSTRACT

Recent years have seen greatly increasing interests in voice over IP in wireless LANs, in which the IEEE 802.11 distributed coordination function protocol or enhanced DCF protocol is used. However, since both DCF and EDCF are contention-based medium access control protocols, it is difficult for them to support the strict QoS requirement for VoIP. Therefore, in this article we propose a novel call admission control scheme that runs at the MAC layer to support VoIP services. The call admission control mechanism regulates voice traffic to efficiently coordinate medium contention among voice sources. The rate control mechanism regulates non-voice traffic to control its impact on the performance of voice traffic. Extensive simulations demonstrate that the proposed schemes can well support statistical QoS guarantees for voice traffic and maintain stable high throughput for non-voice traffic at the same time.

INTRODUCTION

In recent years there has been extensive growth in voice over Internet Protocol (VoIP) in the world. VoIP delivers voice packets over the Internet and hence greatly reduces costs compared to expensive voice calls through the traditional public switched telephone network (PSTN). So far, VoIP is almost limited to the wired part. However, in the near future VoIP is expected to be extended from the Internet to the wireless domain via wireless local access networks (WLANs), due to the diminishing cost of IEEE 802.11 [1] based wireless local access and the increasing deployment of WLANs. Furthermore, some dual-mode cellular and Wi-Fi phone handsets produced recently can switch a cell phone call through cellular networks to a wireless VoIP call through WLANs to reduce the cost. These applications require WLANs to be able to support the strict quality of service (QoS) requirements of voice services. As defined in International Telecommunication Union — Telecommunication Standardization Sector (ITU-T) G.114 [2], for real-time services the tolerable packet loss rate is 1~3 percent, and the one-way transmission delay is preferably shorter than 150 ms but should be no longer than 400 ms.

However, many challenges remain in voice over WLAN (VoWLAN). It is well known that widely deployed IEEE 802.11 WLANs employ a contention-based medium access control (MAC) protocol, the distributed coordination function (DCF). DCF enables fast installation with minimal management and maintenance costs. Although DCF can well support best effort traffic, it may introduce arbitrarily large delay and delay jitters; thus, it is unsuitable for real-time applications with strict QoS requirements. In addition, unlike cellular networks where dedicated channels are assigned to voice traffic, voice packets in WLANs are multiplexed with data traffic. DCF leaves voice streams unprotected. When the best effort traffic load increases, the QoS of VoWLAN could be severely degraded. It is a challenging task to provide QoS for voice traffic while maintaining as high throughput as possible for best effort traffic.

In order to provide QoS for VoWLAN, two QoS mechanisms are necessary: one is call admission control, and the other is distributed channel access control. Call admission control [3–8] normally runs at the access point (AP) and on a call level timescale. The call admission decision is made such that it is feasible (via arbitrary distributed channel access) to provide QoS for all existing voice streams and the newly requested voice stream. Distributed channel access control runs at the mobile stations and on a packet level timescale. Each mobile station needs to contend for the shared channel before transmission. To provide a better QoS level for real-time services, many schemes [9] assign different priorities to real-time traffic and non-real-time traffic. Enhanced Distributed Channel Access (EDCA) defined in IEEE 802.11e [10] supports four access categories (ACs), each of which achieves differentiated channel access by varying the interframe space, and the initial and maximum window sizes for backoff procedures.

1536-1284/06/$20.00 © 2006 IEEE

Although significant research efforts have been made on QoS provisioning in WLANs, two issues are still not well addressed. First, if distributed channel access control is adopted, only service differentiation can be supported; hence, there is no QoS guarantee. Second, most of the existing schemes (including 802.11e EDCA) require the upgrade/replacement of hardware in both APs and mobile stations. Since a lot of DCF-based APs and mobile stations have already been deployed, implementation of such schemes is not practical. Yu *et al.* [11] provided an initial study of this problem and proposed a dual queue strategy, which runs at the MAC layer and does not require modification of the existing hardware. However, it cannot provide QoS guarantee for VoIP flows since best effort traffic is not regulated based on the global traffic condition.

One natural question arises: can we support better QoS than differentiated services for VoIP calls with the widely deployed DCF while achieving high channel utilization? Our previous theoretical studies [12, 13] show that in the IEEE DCF protocol, the strict delay and delay jitter requirements can be statistically guaranteed if the instantaneous aggregate traffic contending for the channel can be controlled below network capacity. Under this condition, the collision probability is small, and the MAC delay is short enough to support voice calls. If the arriving traffic is heavier than the capacity, the WLAN enters saturation, resulting in a significant increase in delay and a decrease in throughput; on the other hand, if the arriving traffic is less than the capacity, channel capacity is wasted. In reality, however, to tune the network to operate at the optimum point requires an effective and efficient control algorithm to regulate input traffic.

In this article we propose a novel call admission and rate control scheme to provide statistical QoS guarantees for VoIP calls. The call admission control mechanism regulates voice traffic to efficiently coordinate the medium contention among voice sources. The rate control mechanism regulates non-voice traffic to control its impact on the performance of voice traffic. One nice feature of our scheme is that it runs on top of the 802.11 MAC protocol; no modification of the firmware of an existing MAC controller chip is needed; it can be considered a software upgrade approach to provide limited QoS for VoWLAN. Extensive simulations demonstrate that the proposed schemes can well support statistical QoS guarantees for voice traffic and maintain high throughput for non-voice traffic at the same time.

A Call Admission and Rate Control Scheme

Overview

In essence, the call admission and rate control (CARC) scheme is responsible for determining when and how packets are passed from the shared outgoing queue to the MAC layer to contend for the shared channel. Therefore, CARC can be considered a control entity lying on top of the MAC sublayer protocol. In this section we first describe the design rationale of CARC and introduce a novel metric, the channel busyness ratio, as the network status indicator to facilitate CARC. Then we present the call admission control and rate control mechanisms, respectively.

Design Rationale and Channel Busyness Ratio

In our previous work [12] we have shown through both theoretical and simulation studies that there is an optimal operating point for IEEE 802.11 DCF. At this point, which corresponds to a certain amount of arriving traffic, the MAC protocol can satisfy the QoS requirements of real-time traffic and achieve maximal channel utilization at the same time. If the arriving traffic is heavier than this threshold, the WLAN quickly enters the saturation state, and the collision probability becomes large, especially when the number of users is large, followed by a significant increase in delay and a decrease in throughput. On the other hand, if the arriving traffic is less than this threshold, the collision probability is small, but channel bandwidth is wasted. Therefore, an effective and efficient control algorithm is required to regulate the input traffic. The algorithm should guarantee that the aggregate traffic rate arriving at the MAC layer is less than or equal to this threshold to provide the required QoS level for voice traffic. It shall also allow the aggregate traffic rate arriving at the MAC layer to be very close or equal to this threshold to maximize the throughput of non-voice traffic (e.g., best effort traffic).

The proposed CARC scheme utilizes the *channel busyness ratio* as the control metric. The channel busyness ratio R_b is the ratio of the time the channel is determined to be busy to the total time. The channel busyness time consists of periods of both successful transmissions and collisions. Let R_s denote the channel utilization, which is the ratio of successful transmission periods to the total time. When the WLAN works at the optimal point, the collision probability is small, $R_b \approx R_s$, and R_b is relatively stable around 0.90 (without request/clear to send, RTS/CTS) or 0.95 (with RTS/CTS) independent of packet size and number of users. Let B_U denote the channel utilization corresponding to the optimal point. CARC should maintain R_b close to B_U to guarantee both a good QoS level and high aggregate throughput.

Figure 1 presents some simulation and analytical results [12] that illustrate the performance of throughput, delay, and delay variation as a function of the channel busyness ratio when RTS/CTS is used. Every node initiates an identical User Datagram Protocol (UDP)/constant bit rate (CBR) traffic flow to a randomly selected neighbor. Different points in Fig. 1 correspond to different sending rates of flows. It can be seen that there is a turning point in all the curves where the channel busyness ratio is about 0.95. Before that point, the throughput keeps increasing; the delay (including queuing delay, channel contention time, and transmission time) and delay variation only slightly increase and are small enough to support real-time traffic. After

The call admission and rate control (CARC) scheme is responsible for determining when and how packets are passed from the shared outgoing queue to the MAC layer to contend for the shared channel. Therefore, CARC can be considered a control entity lying on top of the MAC sublayer protocol.

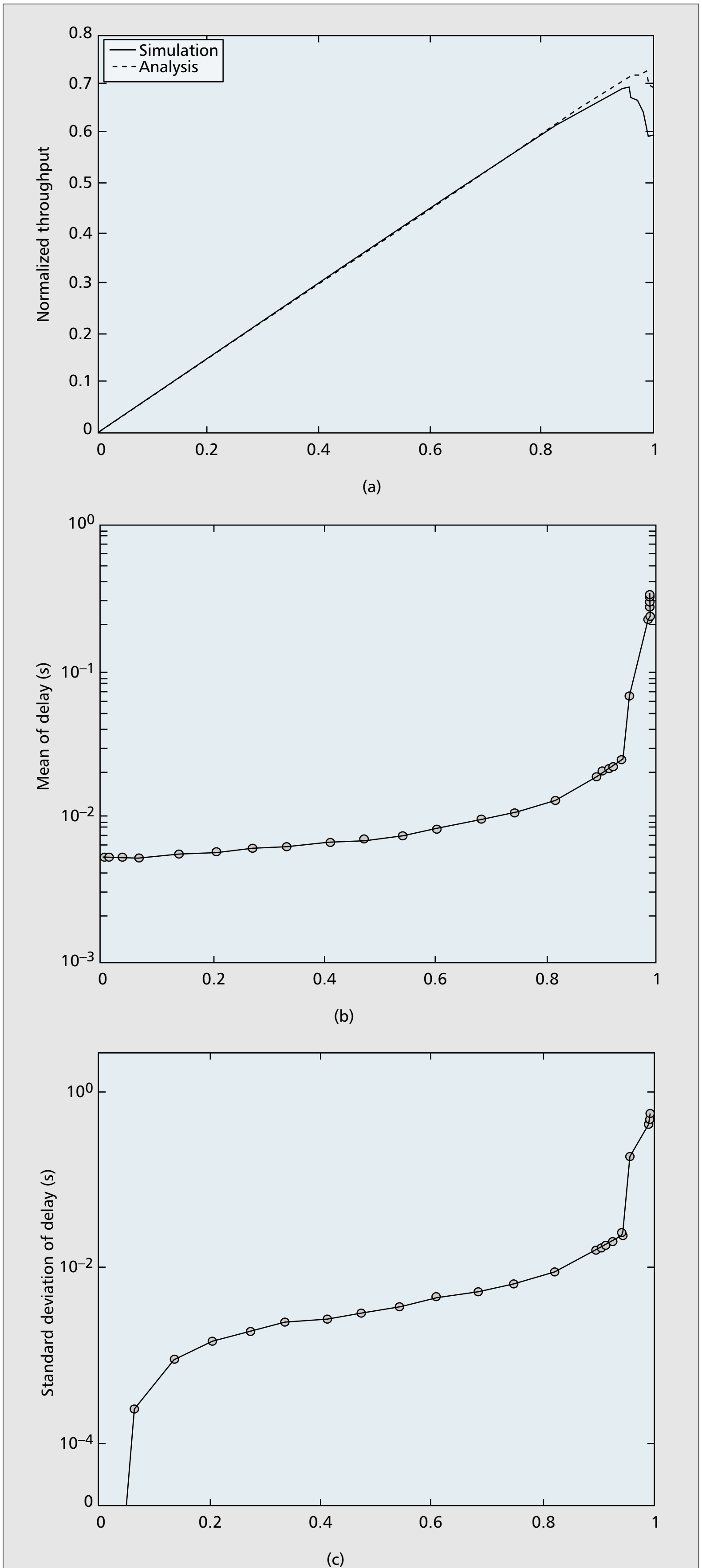


Figure 1. *Throughput and delay performance when there are 50 active nodes and payload size is 8000 bits, channel busyness ratio vs.: a) normalized throughput; b) mean of delay (s); c) standard deviation of delay (s).*

that point, the throughput drops quickly, and the delay and delay variation increase dramatically. Clearly, this turning point is the optimal operating point around which the network should be tuned to operate. The network status is known by keeping track of the channel busyness ratio; hence, it can be used to facilitate the regulation of total input traffic to support QoS.

The channel busyness ratio can also easily be measured in the IEEE 802.11 MAC protocol. There is already a function to determine whether the channel is busy or not at each node. The channel is considered busy whenever the node under consideration is receiving or transmitting, or the network allocation vector (NAV) or physical carrier sensing indicates a busy channel. The channel busyness ratio is equal to the ratio of the sum of all the busy periods to the observation period.

Call Admission Control

The call admission control (CAC) mechanism admits or rejects new voice calls and shall guarantee the QoS level of the admitted calls. In CARC the AP of the WLAN makes the admission decision for each voice call.

CAC admits a new voice call only if the requested resource is available. Here we set an upper bound, denoted B_M, for bandwidth reservation of voice traffic. In our studies we set B_M to 80 percent (it could be adjusted depending on traffic composition) of the maximum channel utilization B_U of the WLAN. This ensures that best effort traffic is operational all the time because best effort traffic is entitled to at least 20 percent of the channel throughput.

The bandwidth requirement of a voice call is characterized by three parameters, (r, r_{peak}, l), where r is the average rate and r_{peak} is the peak rate, both in bits per second, and l is the average packet length in bits. For a call with CBR, $r = r_{peak}$. For a call with variable bit rate (VBR), $r < r_{peak}$. To conduct admission control, we need to convert the traffic rate to the channel utilization u (i.e., the channel time the flow will occupy),

$$u = \frac{r}{l} \times T_{suc}, u_{peak} = \frac{r_{peak}}{l} \times T_{suc}, \quad (1)$$

where T_{suc} is the transmission time of one packet including RTS, CTS, DATA, and acknowledgment (ACK), and all necessary interframe spacings; that is, short interframe space (SIFS) and DCF IFS (DIFS). Thus, (u, u_{peak}) specify a voice flow's bandwidth requirement.

At the AP of the WLAN, the total bandwidth occupied by all admitted voice flows is recorded in two parameters, say the aggregate (u, u_{peak}), denoted by (u_A, u_{peakA}). They are updated when a voice flow joins or leaves through the following admission procedure.

When receiving a voice call request from its application layer, a node must convert the bandwidth requirement into the form of (u, u_{peak}) and send a request with this requirement to the AP, notifying the AP that it wants to establish a voice flow. The AP examines whether there is enough resource to accommodate the new flow. Specifically, it carries out the following procedures:

- If $u_A + u \le B_M$ and $u_{peakA} + u_{peak} \le B_U$, the AP issues a *connection admitted* message, and updates (u_A, u_{peakA}) with ($u_A + u$, $u_{peakA} + u_{peak}$).
- Otherwise, the AP issues a *connection rejected* message to the requesting node.

When a voice flow ends, the source node of the flow should issue a *connection terminated* message to the AP; the latter updates (u_A, u_{peakA}) and responds with a *termination confirmed* message. The AP can also enable a mechanism to release the resource issued to a voice flow when it has no activity for a long time.

To combat the unreliable wireless channel, each node needs to consider the average packet error rate p_e when calculating the bandwidth requirement, using $u/(1 - p_e)$ and $u_{peak}/(1 - p_e)$ instead of the original value of u and u_{peak} in the request message. The factor $(1 - p_e)$ is used because p_e of the entire transmission will fail. Here, we do not consider collision because CARC can well control the number of simultaneously contending nodes at the MAC layer to maintain a very small collision probability.

To ensure short delay of voice packets, they are assigned the highest priority in the outgoing interface queue, which means they are always put at the front of the queue. All control messages related to connection admission and termination are transmitted as best effort traffic. However, they have higher priority than ordinary best effort packets. In this way CARC makes sure that these messages can be transmitted promptly while not affecting the admitted voice flows. Furthermore, when the admitted flows generate more packets than they should, CARC regards these packets as ordinary best effort traffic and will drop them when there is not enough queue space to store them.

Rate Control

The rate control (RC) mechanism regulates the packet sending rate at which the best effort packets are delivered to the MAC layer to contend for the shared channel. RC must ensure that best effort traffic does not affect the QoS level of the admitted real-time traffic but has access to the residual bandwidth left by real-time traffic in order to efficiently utilize the channel.

To satisfy the above requirements, we propose a simple and effective RC scheme for the best effort traffic at each node. Let s ($0 < s < 1$) denote the allowed share of the channel time at one node. It is set as s_{start} initially and is dynamically adjusted by RC. Let t_p be the time a successful transmission of packet p will last over the channel. t_p should include the transmission time of DATA and ACK frames as well as SIFS and DIFS. If RTS/CTS are used, t_p should also include the transmission time of RTS and CTS. With knowledge of s and t_p, RC can calculate the scheduled interval T, the time between two consecutive packets that RC passes to the MAC layer, as

$$T = \frac{t_p}{s}. \tag{2}$$

After the MAC layer finishes transmission of the packet, RC updates s as the currently achievable share, that is, the ratio of t_p to the interval from the time when the MAC layer finishes the last transmission to the current time. To efficiently utilize the channel resource, RC further updates s according to the observed channel busyness ratio measured during the interval between two consecutive packets RC passes to the MAC layer. If the channel busyness ratio is smaller than B_U, it means the channel utilization is small and hence RC should increase s. If the channel busyness ratio is larger than B_U, it means the channel is too busy and hence RC should decrease s to alleviate the chance of collision.

When the channel busyness ratio R_b is less than B_M, the channel is regarded as underloaded in RC. RC adopts a multiplicative-increase law to adjust s, multiplying s by the ratio of B_U to R_b:

$$s \leftarrow s \times \frac{B_U}{R_b}. \tag{3}$$

It is a very aggressive increase law, and the channel busyness ratio will quickly converge to B_U.

When R_b is larger than or equal to B_M and less than B_U, the channel is regarded as moderately loaded in RC. RC adopts an additive-increase law to adjust s, adding s by $t_p/s\delta$:

$$s \leftarrow s + \frac{t_p}{s}\delta, \tag{4}$$

where δ is the increase factor. Notice that t_p/s is the scheduled interval between two consecutive packets passed to the MAC layer. The increase amount of s is proportional to the length of time, and each node will increase its own s by approximately the same amount after the same period. Therefore, RC achieves fair allocation of available channel resource regardless of the current value of s.

When R_b is larger than or equal to B_U, the channel is regarded as heavily loaded. RC adopts a multiplicative-decrease law to adjust s, multiplying s by $\gamma \times B_U/R_b$:

$$s \leftarrow s \times \gamma \times \frac{B_U}{R_b}, \tag{5}$$

where γ is the decrease factor and $0 < \gamma \le 1$. The factor B_U/R_b is to adjust the channel utilization around B_U.

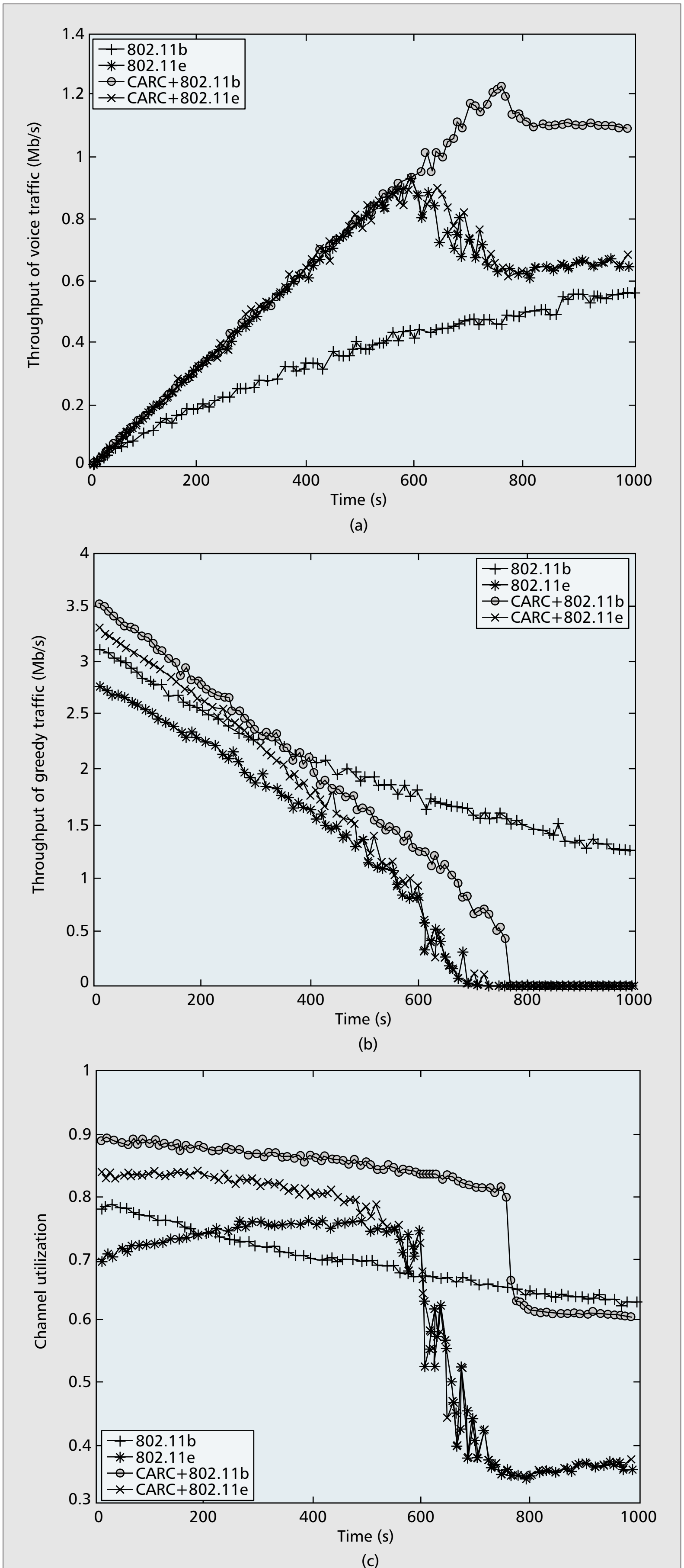
The additive-increase and multiplicative decrease (AIMD) algorithm when $R_b \ge B_M$ can achieve high channel utilization and also good fairness among all the nodes. The values of δ and γ impact the convergence speed of both the efficiency and fairness. In our simulation studies we set $\delta = 0.05$ and $\gamma = 0.95$, and

$$s_{start} = \frac{B_U}{150},$$

which constantly exhibit good performance.

To alleviate packet collision as much as possible and yield channel access opportunity to more urgent real-time traffic, RC also adopts a *packet defer procedure*. At the scheduled packet sending time, the channel may still be very busy because there may already be some nodes concurrently contending for the shared channel at the MAC layer. In this case RC will defer the delivery of the packet to the MAC layer. Specifically, when the scheduled sending time expires, RC checks the observed channel busyness ratio R_b during

In the case where the admitted flows generate more packets than they should, CARC regards these packets as ordinary best effort traffic and will drop them when the queue space is not enough to s tore them.



■ **Figure 2.** *Throughput and channel utilization.*

time period T_s, which starts at the time instant when the last packet is sent to the MAC layer. If it is larger than B_U, an additional delay T_d will be scheduled before passing a packet to the MAC layer, and

$$T_d = T_s (R_b - B_U)/B_U. \quad (6)$$

If R_b is not larger than B_U, RC will immediately pass the packet to the MAC layer.

Before ending this section, we make a few remarks about the proposed RC scheme. In general, the instantaneous traffic rate of real-time services can fluctuate from time to time. When the real-time traffic rate decreases, RC should increase the sending rate of non-real-time or best effort traffic. This is achieved by the multiplicative-increase law when $R_b < B_M$ and the additive-increase law when $B_M \leq R_b < B_U$. When the real-time traffic rate increases, RC should decrease the sending rate of non-real-time or best effort traffic. In this case the real-time packets will occupy the shared channel. The nodes with best effort packets will observe a busy channel and hence defer the packets through the packet defer procedure. This can effectively reduce collisions from best effort traffic and release the channel resource to real-time traffic. Therefore, RC can dynamically adjust the sending rate of non-real-time traffic to accommodate fluctuating real-time traffic, and accordingly effectively control the impact of non-real-time traffic on real-time traffic to provide the required QoS level as well as high channel utilization.

Performance Evaluation

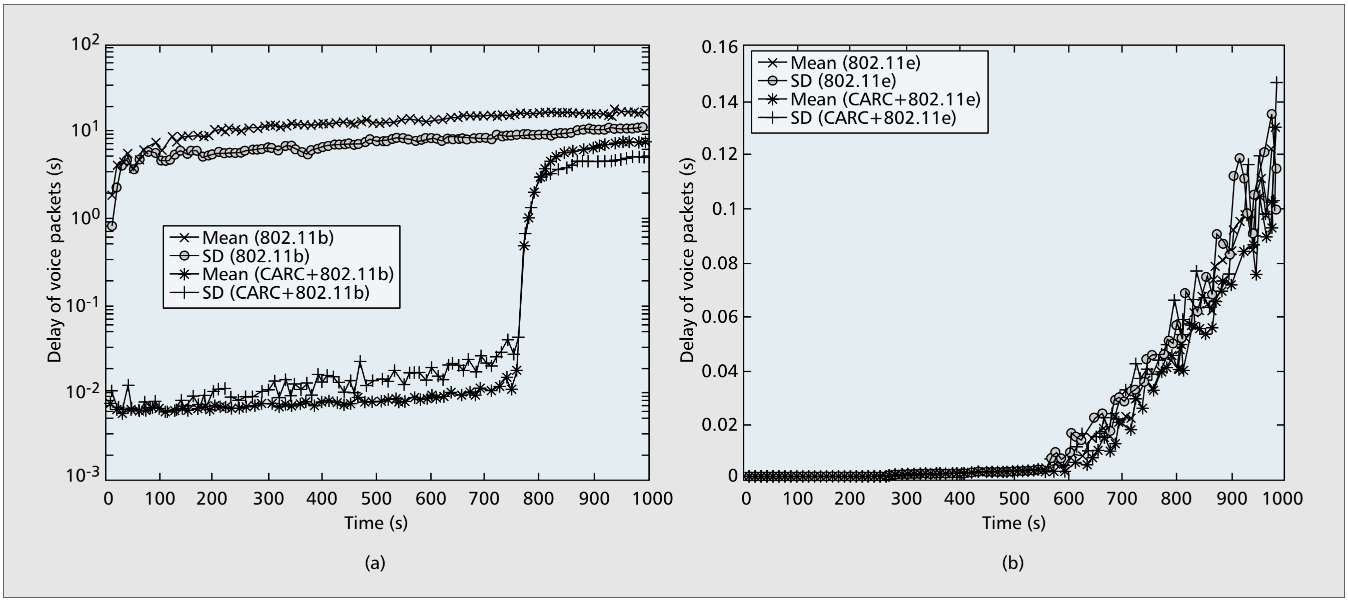
In this section we demonstrate through extensive simulations that CARC can provide QoS guarantees for VoIP flows. The simulation results also show that CARC over 802.11b can support many more VoIP flows than 802.11e and achieve much higher throughput for other traffic as well.

Simulation Setup

To evaluate the performance of the proposed CARC scheme, we conduct simulations in the widely used network simulation tool ns2 (v. 2.27), which we have extended with a CARC module.

In the simulations there are 300 nodes in the WLAN. The parameters of the DCF and enhanced DCF (EDCF) protocols conform to the default settings in the IEEE 802.11b standard and IEEE 802.11e draft [10]. Specifically, in DCF, $DIFS$ = 50 µs, $CW_{\min}$ = 31, and $CW_{\max}$ = 1023. In EDCF, the voice traffic is assigned to AC 3 with the highest priority, and arbitrary distributed IFS (AIFS) $AIFS[3]$ = 50 µs, $CW_{\min}$ = 3, and $CW_{\max}$ = 15; the background traffic is assigned to AC 0 with the lowest priority, and $AIFS[0]$ = 140 µs, $CW_{\min}$ = 15, and $CW_{\max}$ = 1023. In the simulation a number of greedy CBR traffic flows are introduced as background traffic. Each simulation run lasts 300 s. RTS, CTS, and ACK frames are transmitted at a basic rate of 1 Mb/s. DATA frames are transmitted at the channel rate, 11 Mb/s.

We adopt an on/off traffic model for voice traffic. The on and off periods are exponentially



■ **Figure 3.** *Mean and standard deviation (SD) of delay for voice packets.*

distributed with an average value of 300 ms each. During the off periods, there are no voice packets generated. During the on periods, voice packets of each one-way voice flow are generated at a rate of 32 kb/s with a packet size of 160 bytes. For best effort traffic we use the greedy CBR model: there are always best effort packets in the outgoing interface queue. Each best effort packet has a length of 1000 bytes. During the simulation, the RTS/CTS mechanism is used for best effort packets, but not for voice packets because of its short length and relatively large overhead.

In the following simulation we simulate four sets of protocol combinations: 802.11b without CARC, 802.11e without CARC, CARC over 802.11b, and CARC over 802.11e; these are denoted by 802.11b, 802.11e, CARC+802.11b, and CARC+802.11e in the following subsections and figures, respectively.

Performance with a Different Number of Voice Flows

In the first set of simulations there are 50 greedy best effort flows. A new voice flow is periodically added every other 10 s to observe the impact of a new flow on the performance of existing flows. We also disable the admission control part in CARC to find the maximum number of voice flows CARC can support.

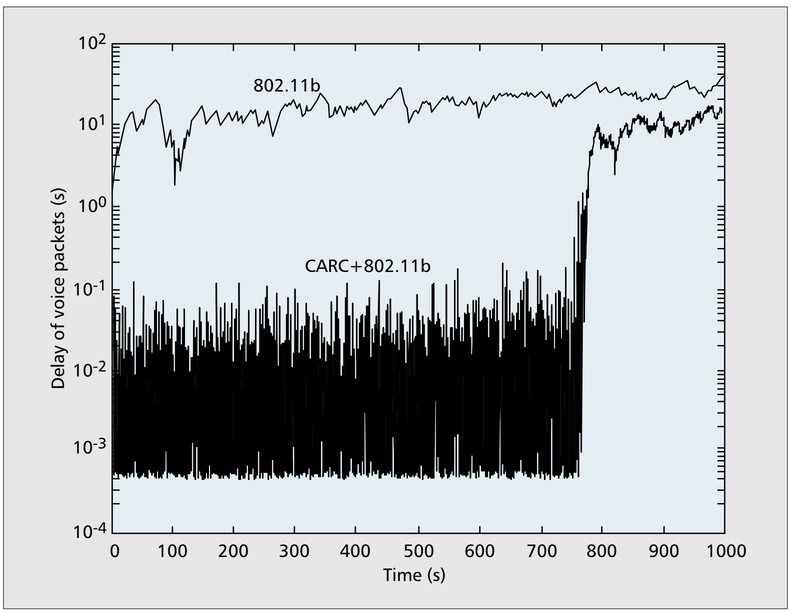
Figure 2a shows the throughput of voice traffic. We can observe that CARC over 802.11b can support 76 voice flows, which is 28.8 percent more than 59 voice flows in 802.11e with and without the CARC scheme. Since 802.11b itself is incapable of controlling the collisions between voice packets and other packets, it cannot support bandwidth requirements for any voice flow when there is some greedy background traffic.

Figure 2b shows the throughput of greedy traffic. It demonstrates that CARC over 802.11b can support much higher throughput for traffic other than VoIP flows and effectively differentiate the voice traffic from other traffic without a differentiated channel access mechanism at the MAC layer. We can observe that 802.11e has the lowest throughput for greedy traffic because greedy traffic has a lower priority to access the channel. CARC can improve the throughput of greedy traffic by up to 20 percent for 802.11e. CARC over 802.11b always has higher throughput for greedy traffic than the previous two cases. When the traffic load of voice flows is higher than the network capacity (i.e., the number of voice flows is larger than 76), the background traffic yields all channel access to voice traffic like 802.11e does, which demonstrates that CARC can also effectively provide higher priority for voice traffic. For 802.11b, since it does not differentiate voice traffic from greedy traffic, it gets through more throughput for greedy traffic than others with a sacrifice of voice throughput when the number of voice flows is larger than 20~30.

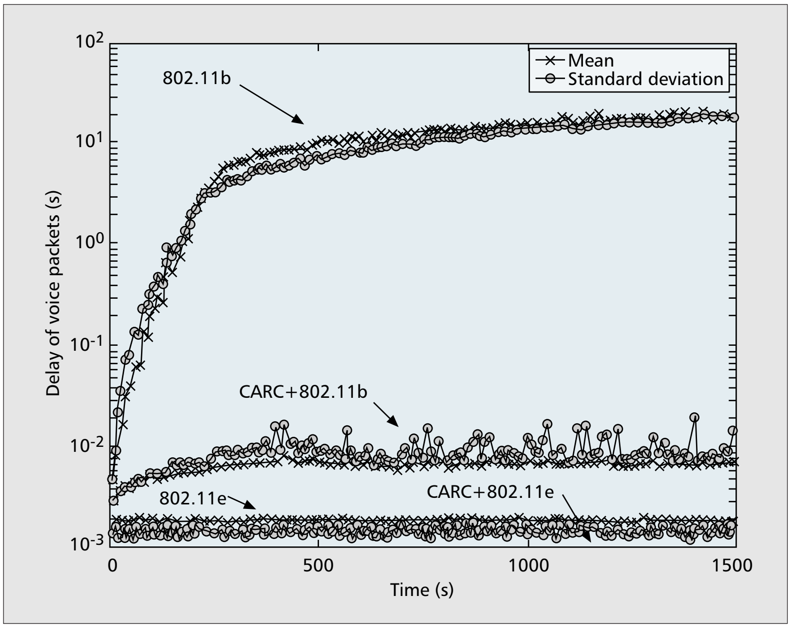
Figure 2c shows the channel utilization. It shows that CARC over 802.11b can be more efficient to utilize the channel resource. Here, channel utilization is calculated by summing up the successful transmission channel time divided by the total channel time. 802.11e has a small start for the channel utilization when the number of voice flows increases. CARC over 802.11e overcomes this problem and has up to 20 percent higher channel utilization than 802.11e itself. CARC over 802.11b has even higher channel utilization than 802.11e with and without CARC, and it has 14 percent higher channel utilization than 802.11b when the number of voice flows is smaller than 76. When this number is larger than 76, the dramatic decrease of channel utilization in CARC over 802.11b is because of the increased channel contention among voice flows.

Now let us study the delay performance. From Fig. 3a, we can observe that 802.11b cannot provide the required delay performance for voice flows because the average value and standard deviation of delay are on a timescale of seconds. In CARC over 802.11b, they ranges from several milliseconds to about 30 ms when there are less than 76 voice flows. Figure 3b shows that 802.11e with and without CARC has similar average delay and delay variation.

Figure 4 shows the delay of all voice packets belonging to the first voice flow. In 802.11b, all packets have a delay in seconds. During 0–760 s in CARC over 802.11b, there are only several packets with a delay larger than 150 ms, no packets with a delay larger than 400 ms, less than 3 percent voice packets with a delay larger than 40.6 ms, and less than 1 percent voice packets with a delay larger than 81.1 ms. From these results, we can conclude that CARC over 802.11b can guarantee the QoS requirements for voice flows. CARC over 802.11b can support many more voice flows than 802.11b, 802.11e, and CARC over 802.11e. Furthermore, CARC over 802.11b also achieves much higher channel utilization than the other cases and allows non-real-time traffic to fully utilize the channel resource left over by real-time traffic.



■ **Figure 4.** *The delay of all packets belonging to the first voice flow.*



■ **Figure 5.** *Delay performance when the number of greedy traffic flows increases along with time.*

If the CAC scheme is enabled in CARC over 802.11b, there will be a total of 52 voice flows that can be admitted in the WLAN. This number can be calculated as follows. Each voice packet consumes 707.27 μs channel time for a successful transmission, consisting of a DATA frame, an ACK frame, a SIFS period, and a DIFS period. During the *on* period of the voice traffic model, the peak rate is 32 kb/s (25 packets/s). By multiplying 707.27 μs by 25, we obtain $(u, u_{peak}) = (0.00884, 0.01769)$ for each voice flow. In the simulation we set $B_U = 0.92$. Thus, the CARC scheme will admit 52 flows because $52 \times 0.01769 = 0.91945$ and rejects incoming voice calls thereafter. These results follow the admission tests from earlier, which have two features: one is to provide QoS guarantee for admitted voice flows; the other is to always allow the best effort traffic to obtain a certain throughput.

Performance with a Different Number of Greedy Traffic Flows

In this set of simulations, there are 30 voice flows at the beginning, and one greedy CBR source node is added every other 10 s.

Figure 5 illustrates that delay in 802.11b quickly increases along with the number of greedy traffic flows while in 802.11e, CARC over 802.11b and CARC over 802.11e, both mean and standard deviation of delay are around several milliseconds independent of the number of greedy traffic flows. We also observe from the simulation results that, in 802.11b, there are a lot of packets with a delay larger than 400 ms when there are more than three greedy traffic flows in the WLAN. That is to say, 802.11b itself cannot support QoS requirements of voice flows when there are four or more greedy traffic flows. On the other hand, in 802.11e CARC over 802.11b and CARC over 802.11e, there are almost no packets with a delay larger than 400 ms during the whole 1500 s.

The Impact of Channel Errors

To examine how the proposed CARC scheme performs when there exist channel errors, we conduct a set of simulations with different packet error rates to observe the maximum number of voice flows the protocol can support. Here, two requirements must be satisfied to determine whether a certain number of voice flows can be supported: all successful transmitted voice packets have a delay less than 400 ms, and the packet loss rate is less than 1 percent. We examine the packet loss rate for each 10 s interval to check whether it is satisfied or not. In the simulation one voice flow is added every other 10 s. The maximum number of voice flows that can be supported is obtained by counting the number of existing voice flows until adding one more voice flow results in the violation of the 400 ms delay limit or the requirement of packet loss rate.

The simulation results illustrate that CARC still support QoS level for voice flows when the channel error rate is less than 0.50. The higher the packet error rate, the less the number of supported voice flows. From Fig. 6, we can observe that CARC over 802.11b can support up to 39.4 percent more voice flows than 802.11e does. When the channel error rate is equal to 0.50, seven retransmissions for the voice packets at the MAC layer is not enough to satisfy that the packet loss rate is less than 1 percent during all 10 s periods. This is because that 0.5^7 = 0.78125 percent, which is close to 1 percent, and there is packet collision as well.

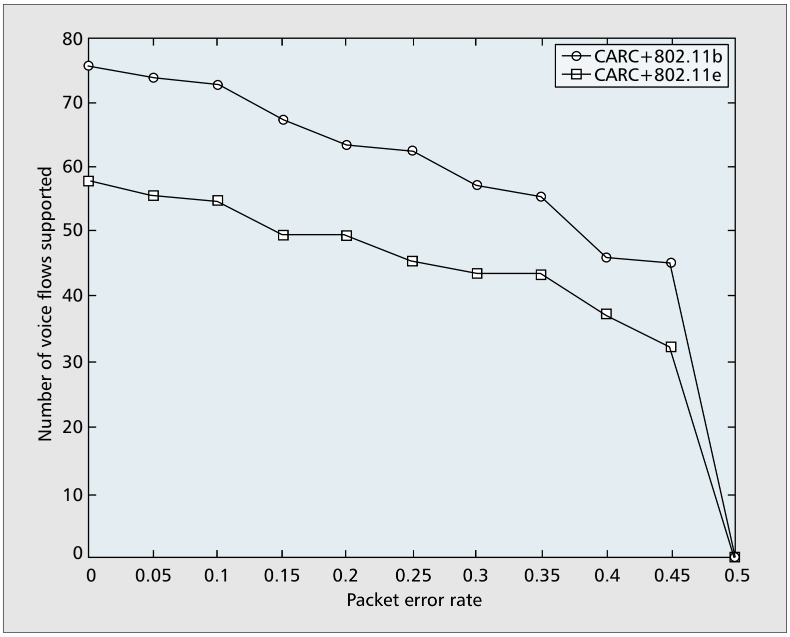


Figure 6. *Maximum number of voice flows supported in CARC.*

Conclusion

In this article we propose a novel call admission and rate control protocol to support statistical QoS guarantee for voice over IP traffic in IEEE 802.11 wireless LANs. Based on the novel use of the channel busyness ratio, the proposed protocol can statistically guarantee stringent QoS requirements of voice over IP traffic while achieving high channel utilization.

Acknowledgment

This work was supported in part by the National Science Foundation under Faculty Early Career Development Award ANI-0093241 and under grant DBI-0529012, and by the Officer of Naval Research under grant N000140210464 (ONR Young Investigator Award).

References

[1] "IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," ISO/IEC 8802-11: 1999(E), Aug. 1999.
[2] ITU-T G.114, "One-Way Transmission Time," 1996.
[3] X. Chen, H. Zhai, and Y. Fang, "Enhancing the IEEE 802.11e in QoS Support: Analysis and Mechanisms," *Proc. 2nd Int'l. Conf. QoS in Heterogeneous Wired/Wireless Networks (QShine'05)*, Orlando, FL, Aug. 2005.
[4] X. Chen *et al.*, "Supporting QoS in IEEE 802.11e Wireless LANs," to appear, *IEEE Trans. Wireless Commun.*
[5] D. Gao, J. Cai, and K. N. Ngan, "Admission Control in IEEE 802.11e Wireless LANs," *IEEE Network.*, Special Issue on Wireless Local Area Networking: QoS Provision and Resource Management, vol. 19, no. 4, July/Aug. 2005, pp. 6–13.
[6] Y. Xiao and H. Li, "Evaluation of Distributed Admission Control for the IEEE 802.11e EDCA," *IEEE Commun. Mag.*, vol. 42, no. 9, Sept. 2004, pp. S20–S24.
[7] H. Zhai, X. Chen, and Y. Fang, "A Call Admission and Rate Control Scheme for Multimedia Support over IEEE 802.11 Wireless LANs," *1st Int'l. Conf. QoS in Heterogeneous Wired/Wireless Networks (QShine'04)*, Dallas, TX, Oct. 2004.
[8] H. Zhai, X. Chen, and Y. Fang, "A Call Admission and Rate Control Scheme for Multimedia Support over IEEE 802.11 Wireless LANs," to appear, *ACM Wireless Networks*.
[9] W. Pattara-Atikom, P. Krishnamurthy, and S. Banerjee, "Distributed Mechanisms for Quality of Service in Wireless LANs," *IEEE Wireless Commun.*, vol. 10, no. 4, June 2003, pp. 26–34.
[10] "IEEE Standard for Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS)," IEEE 802.11e/D13.0, July 2005.
[11] J. Yu, S. Choi, and J. Lee, "Enhancement of VoIP over IEEE 802.11 WLAN via Dual Queue Strategy," *Proc. IEEE ICC '04*, Paris, France, June 2004.
[12] H. Zhai, X. Chen, and Y. Fang, "How Well Can the IEEE 802.11 Wireless LAN Support Quality of Service?" *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, Nov. 2005, pp. 3084–94.
[13] H. Zhai, Y. Kwon, and Y. Fang, "Performance Analysis of IEEE 802.11 MAC Protocols in Wireless LANs," *Wireless Commun. and Mobile Comp.*, vol. 4, Dec. 2004, pp. 917–31.

Biographies

HONGQIANG ZHAI [S'03] (zhai@ecel.ufl.edu) received B.E. and M.E. degrees in electrical engineering from Tsinghua University, Beijing, China, in July 1999 and January 2002, respectively, and is currently working toward a Ph.D. degree at the Department of Electrical and Computer Engineering, University of Florida, Gainesville.

JIANFENG WANG (jfwang@ufl.edu) [S] received the B.E. and M.E. degrees in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, in 1999 and 2002, respectively. Now he is working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Florida.

YUGUANG FANG [S'92, M'94, S'96, M'97, SM'99] (fang@ece.ufl.edu) received a Ph.D. degree in systems engineering from Case Western Reserve University in January 1994 and a Ph.D. degree in electrical engineering from Boston University in May 1997. He was an assistant professor in the Department of Electrical and Computer Engineering at New Jersey Institute of Technology from July 1998 to May 2000. He then joined the Department of Electrical and Computer Engineering at University of Florida in May 2000 as an assistant professor, got an early promotion an associate professor with tenure in August 2003 and to professor in August 2005. He has published over 150 papers in refereed professional journals and conferences. He received the National Science Foundation Faculty Early Career Award in 2001 and the Office of Naval Research Young Investigator Award in 2002. He has served on many editorial boards of technical journals including *IEEE Transactions on Communications*, *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Mobile Computing*, and *ACM Wireless Networks*.