Capacity Improvement and Analysis for Voice/Data Traffic over WLANs

Ping Wang, Hai Jiang, *Member, IEEE,* and Weihua Zhuang, *Senior Member, IEEE*

***Abstract*— Voice over wireless local area network (VoWLAN) is an emerging application taking advantage of the promising voice over Internet Protocol (VoIP) technology and the wide deployment of WLANs all over the world. The real-time nature of voice traffic determines that controlled access rather than random access should be adopted. Further, to fully exploit the capacity of the WLAN supporting voice traffic, it is essential to explore statistical multiplexing and to suppress the large overhead. In this paper, we propose mechanisms to enhance the WLAN with voice quality of service (QoS) provisioning capability when supporting hybrid voice/data traffic. Voice multiplexing is achieved by a polling mechanism in the contention-free period and a deterministic priority access for voice traffic in the contention period. Header overhead for voice traffic is also reduced significantly. Delay-tolerant data traffic is guaranteed an average portion of service time in the long run. A session admission control algorithm is presented to admit voice traffic into the system with QoS guarantee. Analytical and simulation results demonstrate the effectiveness and efficiency of our proposed solutions.**

***Index Terms*— Voice over Internet Protocol (VoIP), wireless local area network (WLAN), voice over WLAN, quality of service (QoS), capacity, session admission control.**

I. INTRODUCTION

THE convergence of the voice over Internet protocol (VoIP) technology and the wide deployment of wireless local area networks (WLANs) has driven the application of voice over WLAN (VoWLAN), which is expected to experience a dramatic increase in the near future. Fig. 1 shows a typical scenario when the traffic of voice conversations and data transfers passes through the access point (AP). For voice service, at the sender, the analog voice signal is compressed and encoded by a codec. After the inclusion of the RTP (real-time transport protocol)/UDP (user datagram protocol)/IP header during the packetization procedure at the transport and network layers, the voice packets are transmitted over the networks and finally to the receiver end. At the receiver, a playout buffer is usually used to alleviate the effect of delay jitter. Then the receiver applies de-packetization and decoding to recover the original voice signal.

Manuscript received August 15, 2005; revised December 13, 2005; accepted December 20, 2005. The associate editor coordinating the review of this paper and approving it for publication was Q. Zhang. This work was supported by a research grant from the Natural Science and Engineering Research Council (NSERC) of Canada.

P. Wang and W. Zhuang are with Department of Electrical and Computer Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1 (e-mail: {p5wang, wzhuang}@bbcr.uwaterloo.ca).

H. Jiang is with the Electrical Engineering Department, Princeton University, Princeton, New Jersey 08544, U.S.A. (email: haijiang@princeton.edu).

Digital Object Identifier 10.1109/TWC.2007.05630.

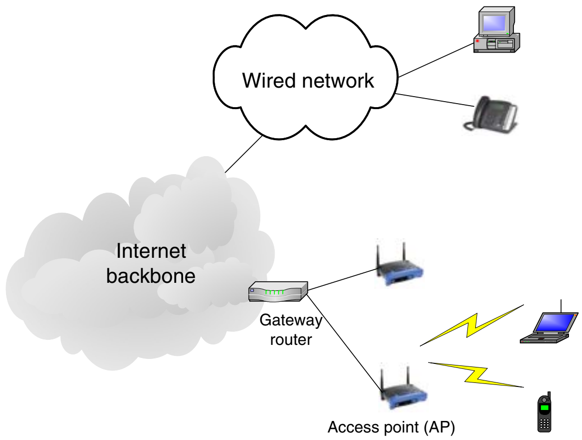


Fig. 1. The architecture for WLANs supporting voice and data services.

There exist two major challenges for VoWLAN. One challenge is how to increase the system capacity for voice users. It has been found out that the system capacity for voice users is quite low in current WLANs, far from what is needed [1]–[3]. Originally designed for data traffic, the WLANs experience bandwidth inefficiency when supporting voice traffic due to the large overhead. Hence, it is essential to enlarge the VoIP capacity supported by WLANs. The other challenge is quality of service (QoS) provisioning for voice users. Voice traffic is sensitive to delay and delay jitter. In current WLANs, VoIP traffic may be interfered by other traffic (e.g., data traffic), resulting in a delay bound violation or large delay variance. Therefore, it is necessary to enhance QoS support capability over WLANs.

The focus of this paper is to address these two challenges. We base our work on IEEE 802.11e since it is the most promising technology for QoS provisioning in WLANs. With minor modifications to IEEE 802.11e, we can increase the system capacity significantly for voice traffic, provide guaranteed QoS to voice users, and provide data traffic a certain level of service share. Specifically, the contributions of this paper are as follows:

- We propose an efficient resource allocation scheme which combines controlled access with contention based access to achieve voice traffic multiplexing. During the contention period, a deterministic priority access scheme is proposed with a minor modification to IEEE 802.11e in order to provide guaranteed QoS to voice traffic. Delay-tolerant data traffic is guaranteed an average portion of service time.
- We increase the system capacity significantly in terms of

voice session number by effective reduction of the system overhead.

- We investigate the performance of voice service in both contention period and contention-free period, analyze voice capacity by considering voice traffic multiplexing and packet loss requirement, and provide a session admission control algorithm to guarantee the QoS of voice and data.

The rest of this paper is organized as follows. In Section II, background and related work are discussed. Section III presents the proposed capacity improvement mechanisms. In Section IV, we provide the performance analysis and present a session admission control algorithm. Numerical results and discussion are given in Section V, followed by conclusion remarks in Section VI. As many symbols are used in this paper, Table I summarizes the important ones.

II. Background

A. Limitations of IEEE 802.11 in Supporting Voice

As a real-time application, VoWLAN is delay-sensitive but can tolerate a certain level of packet loss. Each voice packet should be transmitted within a delay bound. Also, the delay jitter (i.e., variation of voice packet delay) should be carefully controlled as it may degrade voice quality more severely than delay. Traditionally, an appropriately designed playout buffer is an effective way to deal with delay jitter and make the voice understandable. Therefore, delay bound and packet loss rate guarantees are the main QoS requirements for voice service under consideration in this paper.

As the most popular WLAN standard, IEEE 802.11 defines a mandatory distributed coordination function (DCF) and an optional centralized point coordination function (PCF). DCF is based on the carrier sense multiple access with collision avoidance (CSMA/CA), where the collision is resolved by binary exponential backoff. The optional request-to-send (RTS)/clear-to-send (CTS) dialogue can also be applied to further deal with the hidden terminal problem. Mainly designed for data transmission, DCF does not take into account the delay-sensitive nature of real-time services. On the other hand, with PCF, a contention-free period (CFP) and a contention period (CP) alternate periodically. During a CFP, when polled, a station gets the permission to transmit its frames. The main drawbacks of PCF include bandwidth waste when two stations in the same basic service set (BSS) (which is composed of an AP and a number of stations associated with the AP) try to communicate with each other, uncontrolled transmission time of polled stations, and unpredictable CFP start time [4].

To enhance the legacy IEEE 802.11 medium access control (MAC), IEEE 802.11e proposes new features with QoS provisioning to real-time applications. As an extension of DCF, the enhanced distributed channel access (EDCA) provides a priority scheme to distinguish different traffic categories by classifying the arbitration interframe space (AIFS), and the initial ($CW_{\min}$) and maximum ($CW_{\max}$) contention window sizes in the backoff procedures. In IEEE 802.11e, the hybrid coordination function (HCF) can assign specific transmission durations by a polling mechanism. A station can be polled in either CFP or CP. In addition, the direct link protocol allows a station to transmit frames directly to another station.

DCF/EDCA is not effective or efficient in supporting the delay-sensitive voice traffic. The contention-based nature and exponential backoff mechanism cannot guarantee that a voice packet is successfully delivered within the delay bound. In addition, the time to transmit the payload of a voice packet is only a very small portion of the total time to transmit the packet, due to the overhead such as the RTP/UDP/IP headers, MAC header, physical (PHY) preamble, the IFSs, and the backoff time. Consequently, the capacity to accommodate voice traffic in DCF or EDCA is very limited. For example, IEEE 802.11b can support approximately 11 simultaneous two-way voice sessions if a GSM 6.10 codec is used [1].

In order to guarantee the delay requirement of voice service, controlled access is preferred in the WLAN, in which the AP polls each voice station periodically. To efficiently utilize the radio resources, two challenging issues need to be addressed:

- Voice multiplexing – Generally, voice traffic can be represented by an `on/off` model: a voice user is alternately in talk spurt (`on` state) and in silence (`off` state). The durations of the `on` and `off` states are independently and exponentially distributed with parameter α and β, respectively. At an `on` state, voice packets are generated periodically with an inter-arrival time t_a, while no voice packet is generated at an `off` state. It is desired to achieve the statistical multiplexing in VoWLAN based on this property.
- Overhead suppression – The overhead due to the large header may significantly degrade the system efficiency, and should be suppressed as much as possible.

B. Related work

In recent years, VoWLAN has drawn a lot of attention from the R&D community. To provide priority to real-time traffic, EDCA is defined in IEEE 802.11e. It applies different initial and maximum contention window sizes and different IFS values to provide differentiation to different types of traffic. However, it provides only statistically rather than deterministically prioritized access to high priority traffic such as real-time voice. In other words, the prioritized access for high priority traffic is only guaranteed in a long term, but not for every contention. Since each station continues to count down its backoff timer once the channel becomes idle for an IFS, a low priority packet with a probably large initial backoff timer will eventually count down its backoff timer to a small value, most likely smaller than the backoff timer of a newly backlogged high priority packet. Then the low priority packet grabs the channel, resulting in the high priority packet waiting for a long time for the next competition [5]. Such statistically prioritized access is hard to satisfy the delay requirement of each voice packet. Furthermore, when applying EDCA, with the increase of low priority traffic loads, the collision probability seen by the high priority traffic increases. High priority traffic can suffer performance degradation due to low priority traffic offering heavy loads [6].

Many efforts have been made on voice traffic capacity analysis. Both experimental results [3] and analytical results

TABLE I
SUMMARY OF IMPORTANT SYMBOLS USED.

Symbol	Definition
B	Average burst size
$CW_{\min}$ ($CW_{\max}$)	Initial (maximum) contention window size
L	Payload size of a voice packet
M	The maximum number of downlink (or uplink) voice packets generated in a service interval from a voice session
N	Voice session number in a service interval
n	Number of contending voice sessions in a CP
n_1 (n_2)	Number of contending voice sessions with contention window CW_1 (CW_2) in a CP
$P_d(i)$ ($P_u(i)$)	The probability of generating i downlink (uplink) voice packets from a voice session for transmission in a CFP
P_L	Packet loss rate bound in the CFP
$p_{(n_1,n_2;\ l_1,l_2)}$	The transition probability from state (n_1, n_2) when $l_1(\leq n_1)$ and $l_2(\leq n_2)$ transmissions are from voice sessions with contention window CW_1 and CW_2, respectively
R	Voice payload transmission rate
$s_{(n_1,n_2;\ l_1,l_2)}$	The next state if the current one is (n_1, n_2) and the number of transmissions from voice sessions with contention window CW_1 and CW_2 are $l_1(\leq n_1)$ and $l_2(\leq n_2)$, respectively
T_{CFP} (T_{CP})	The duration of the CFP (CP) in a service interval
$T_{CFP}^m(N)$	The minimum T_{CFP} needed in order to guarantee the packet loss rate bound of voice if N voice sessions are admitted.
$\overline{T_{CP}^v}$ ($\overline{T_{CP}^v}(N)$)	Average time in a CP used to serve contending voice sessions (with N admitted voice sessions)
T_{SI}	Service interval
$T(n_1, n_2)$	Average time needed for transitions from state (n_1, n_2) to the absorbing state (0, 0)
t_a	Voice packet inter-arrival time
$t_{(n_1,n_2;\ l_1,l_2)}$	The transition time from state (n_1, n_2) when $l_1(\leq n_1)$ and $l_2(\leq n_2)$ transmissions are from voice sessions with contention window CW_1 and CW_2, respectively
$1/\alpha$ ($1/\beta$)	Mean `on` (`off`) period of voice
τ	Slot duration
ϕ	The average fraction of time required for data service
$\xi(N)$	The outage probability that T_{CP} is not sufficient to serve all the contending voice sessions when N voice sessions are admitted

[2] have demonstrated that system capacity for voice traffic is very limited in WLANs due to the large header overhead and the inefficiency of IEEE 802.11 MAC protocol. Most of previous work assumes that the voice traffic is constant rate traffic, which is not the case in reality. For more accurate capacity estimate, the `on/off` model should be applied and the voice traffic multiplexing should be considered. Further, all the above work focuses on DCF based (or contention based) WLANs. However, as mentioned, DCF is not effective or efficient to support the delay-sensitive voice traffic. Controlled access is more suitable for voice traffic delivery because of its less overhead and guaranteed delay performance. Unfortunately, only very limited work focuses on controlled access. The capacity of PCF is analyzed in [7], [8]. To the best of our knowledge, there is no analysis of HCF so far.

In order to improve the capacity of voice traffic over WLANs, various solutions have been proposed [1], [9]–[12]. A cyclic shift and station removal polling scheme is proposed in [9] to take advantage of the multiplexing of voice packets. Without changing the IEEE 802.11 MAC protocol, VoIP capacity is increased by reducing the header overhead of voice packets in [1], [10]. A voice multiplex-multicast (M-M) scheme is proposed in [1], in which the AP multiplexes packets from several VoIP streams into one multicast packet for transmission. However, an additional delay is expected in composing such a composite packet. In [10], compressed RTP is used to reduce the VoIP header; however, the overhead incurred by IP, MAC and physical layer remains high. On the other hand, some research [11], [12] increases the capacity by introducing new MAC protocols. By reducing the number of collisions or reducing the idle time caused by backoff, these new MAC protocols achieve a better throughput than the IEEE 802.11 MAC protocol, resulting in an increased capacity. These MAC protocols are all contention based, and none of the studies focus on controlled access.

Since the system capacity is very limited in WLANs, ses-

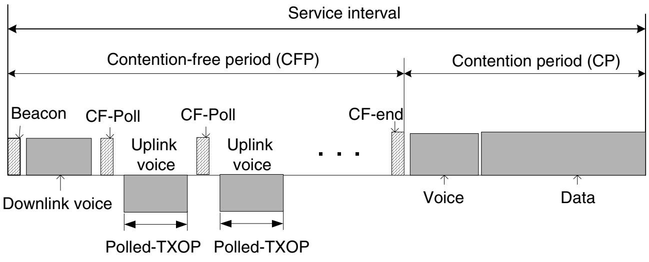


Fig. 2. The structure of a service interval.

sion admission control is important and necessary to maintain the QoS of existing sessions. As revealed in [3], an additional session that exceeds the system capacity will cause unacceptable quality for all ongoing sessions. Previous research on session admission control in WLANs can be classified into two categories. One is analysis based and the other is measurement based. Analysis based admission control algorithms, including [13], [14] and the reference model provided in IEEE 802.11e, make admission decision based on the knowledge of the system capacity derived from analysis. The measurement based algorithms [10], [15], [16] make admission decisions based on the measurement or estimate of channel utilization. Based on the measurements of the fraction of time per time unit needed to transmit the flow over the network [15], collision statistics of each flow [16], or the transmission time of each traffic type [10], available/residual budgets are calculated for admission control.

III. Proposed Mechanisms for Capacity Improvement

We use the same structure of the IEEE 802.11e HCF service interval, and the beacon interval is equal to the service interval. Voice and data services are supported. In each service interval, there are two periods: CFP and CP, as shown in Fig. 2. The CFP is used to accommodate voice stations in the downlink (from the AP to the mobile stations) and uplink (from mobile stations to the AP) by polling. For the uplink transmission, the AP sends a CF-Poll frame which grants each polled station a transmission opportunity (TXOP). No acknowledgement (ACK)/retransmission is required for voice transmission in order to avoid the retransmission delay. In the CP, the AP and all the stations can contend for the channel. It is mainly used to serve data stations and to transmit the first few packets of each voice station's uplink talk spurts. To guarantee the priority of voice over data in the CP, voice packets are always transmitted ahead of data packets, to be discussed in Section III-B. The length of a service interval is fixed and depends on the delay bound of voice traffic. The length of the CFP and CP depend on the voice and data traffic load and QoS provisioning technique.

The QoS enhancement in our proposed scheme consists of three mechanisms: voice traffic multiplexing, deterministic access priority of voice, and overhead reduction, as elaborated in the following.

A. Voice Traffic Multiplexing

In order to achieve a high resource utilization, the network designers should consider the `on/off` characteristic of voice traffic, so that resources are allocated to stations only when they are in a talk spurt. However, IEEE 802.11e does not describe a polling method in HCF to achieve voice traffic multiplexing. Generally, it is easy for the AP to recognize the ending moment of a talk spurt, but it is difficult to know the exact starting moment of a talk spurt. The AP may still need to poll a voice station even during its silent period in order not to miss the beginning of a talk spurt, which is not efficient considering the polling overhead. Here we propose a more efficient polling mechanism to achieve the voice traffic multiplexing.

Consider the case when a station initiates a voice session. If the session can be admitted, the AP will add the station to its polling list. Since the duration of each service interval (T_{SI}) is fixed and the voice packet inter-arrival time (t_a) is a constant in a talk spurt, each station (in the `on` state) will be granted a fixed TXOP just enough to accommodate the generated voice packets during a service interval. If a polled station has no packet to send or cannot use up all the time of TXOP, the AP considers the station being in the silent period and deletes it from its polling list, except the newly added (to the polling list) stations. When a previously `off` station has voice packets to send, the station will contend for the channel during the next CP. Once it gets the channel, it will send out all the voice packets in the buffer (as long as the transmission time does not exceed the TXOP). The AP monitors all the packets transmitted in each CP. For every voice packet, the AP records the sender address (or ID) and adds it to the polling list. If the station is newly added to the list during the last service interval, the AP will retain it in the list, even though it may not use up all the TXOP or has no packet to send in the current CFP, since a few voice packets at the beginning of a talk spurt were sent during the last CP.

Once a voice station is added to the polling list, all the subsequent voice packets in the same talk spurt will be transmitted in the CFP. Hence, the voice station does not need to contend for the channel anymore for the current talk spurt.

B. Deterministic Access Priority of Voice over Data in the CP

Another challenging issue is raised from the uplink voice multiplexing: to meet the strict delay requirement of uplink voice traffic, it should be guaranteed that a voice station can access the channel successfully during the CP when needed.

To provide QoS guarantee for voice traffic regardless of the data traffic load in the WLAN, data stations should not transmit in the CP until no voice station contends for the channel. As discussed in Section II-B, EDCA cannot meet this requirement. As a result, a deterministically prioritized access for voice traffic is more appropriate. Only a few voice packets at the beginning of each talk spurt need to contend in a CP, which should not significantly degrade the QoS of data traffic if a deterministically prioritized access is provided to voice.

A simple way to provide deterministically prioritized access is to modify EDCA so that the AIFS of data access category (AC) (AIFS[AC_data]) is equal to (AIFS[AC_voice]+$CW_{\max}$[AC_voice]), the summation of AIFS of voice AC and the maximum contention window size of voice AC. However, it is not efficient in terms of channel utilization. The number of contending voice packets is expected to be small in a CP, and all the data packets have to wait a long time before getting the channel, resulting in a waste of resources.

Inspired by the idea of black-burst contention [17], here we propose an efficient mechanism to provide deterministically prioritized access to voice, by minor modifications to IEEE 802.11e EDCA. In our mechanism, the AIFSs for voice traffic and data traffic remain the same as those in EDCA. In addition, the contention behaviors for data stations remain the same as in EDCA. The contention behaviors of voice stations are modified as follows. For a contending voice station, after waiting for the channel to be idle for AIFS[AC_voice], instead of further waiting for the channel to be idle for a duration of backoff time, the voice station will send a busy tone, and the length of the busy tone (in the unit of slot time) is equal to its backoff timer. After the completion of its own busy tone, the station monitors the channel for the duration of a slot time. If the channel is still busy (which means that at least one other voice station is sending busy tone), the station will quit the current contention, keep its contention window, choose a backoff timer randomly from its contention window, and wait for the channel to be idle for AIFS[AC_voice] again. Otherwise, the station (which sends the longest busy tone) will send its voice packets. It is possible that two or more voice stations happen to send the same longest busy tone, resulting in a collision. Contention windows of collided stations evolve by the same way as that in EDCA, and each collided voice station chooses a backoff timer randomly from its contention window for the next contention. Since there is no ACK frame sent back to acknowledge the successful voice transmission, it is difficult for the sender to recognize the collision. To address the problem in our scheme, for the first packet from a voice station received in a CP, the receiver should send back an ACK frame to the sender. The voice sender continues to contend in the CP if no ACK is received.

In a CP, if there exists at least one voice contender, all data stations will sense the busy tone during the AIFS[AC_data] (>AIFS[AC_voice]), and defer their transmissions. When a collision happens between voice stations, the data stations will wait for the channel to be idle for the duration of ACK timeout plus AIFS[AC_data] before they attempt to acquire the channel, which ensures that voice stations will not lose the channel access priority to the data stations even when a collision happens. Furthermore, when all the active (in terms of uplink transmission) voice stations are included in the polling list, the data stations can make full use of the CP resources.

By using the above mechanism, it seems that the waiting time (before getting the channel) of a voice station is larger than that in EDCA, since the voice station with the largest backoff timer instead of the smallest backoff timer (as in EDCA) gets the channel. However, as the number of voice stations contending for the channel simultaneously is very likely to be small, the initial and maximum window sizes for voice AC can be set to small values, so the negative effect of waiting time should be negligible in our mechanism.

C. Voice Overhead Reduction

To support voice over WLANs, it is important to reduce the overhead and improve the transmission efficiency over the radio link. The large packet header overhead can significantly affect the capacity of the WLAN in supporting voice service. For example, if a GSM 6.10 codec is used, a voice packet payload is 33 bytes while the RTP/UDP/IP headers are 40 bytes. In addition, the PHY preamble, MAC header, and control packets all consume bandwidth. As a result, the overall efficiency is less than 3% [1]. Actions need to be taken to alleviate the effect of the overhead.

Recently, various header compression techniques for VoIP have been proposed. The RTP/UDP/IP headers can be compressed to as small as 2 bytes [18], [19]. The compression technique is adopted in our research.

In our proposed scheme, the PHY and MAC layer overheads are further reduced by aggregating the buffered voice packets from or to a voice station together and transmitting them by one MAC frame. Take uplink transmissions as an example. The AP polls each voice station periodically after every service interval, which depends on the delay bound of voice traffic. Within each service interval, several voice packets may be generated and buffered by each voice station. In order to increase the efficiency, we combine the payload of these packets together and add a common MAC layer header instead of sending them one by one. It reduces the overall MAC layer header and PHY preamble overhead.

IV. Performance Analysis

For a WLAN supporting voice/data traffic, we assign a higher priority to voice traffic. The CFP is used to transmit voice traffic; and in the CP, voice traffic has deterministic priority over data traffic. To provide data traffic a certain level of QoS, it is required that the average service time in each CP for data traffic is at least a pre-specified fraction (ϕ) of the whole service interval. Hence, we need to determine the maximum number of voice sessions[1] that can be supported by the average fraction (1-ϕ) of time in each service interval.

A. Average Time Required to Serve Contending Voice Sessions in a CP

In each CP, if there is any contending voice session, the whole CP time can be partitioned to two portions: the first portion is used by voice sessions to contend and transmit, while the second portion is for data traffic. In the following, the average time needed to serve the contending voice sessions in a CP is derived.

In a CP, consider n voice sessions contending for the channel. For simplicity of presentation, the contention window of each voice session takes values from the set $\{CW_1, CW_2\}$ where $CW_2 = 2 \cdot (CW_1 + 1) - 1$, and at the beginning of each CP, all contending voice sessions are with CW_1. Our analysis can be easily extended to cases with 3 or more choices

[1] In this paper, a "voice session" means a two-way voice session.

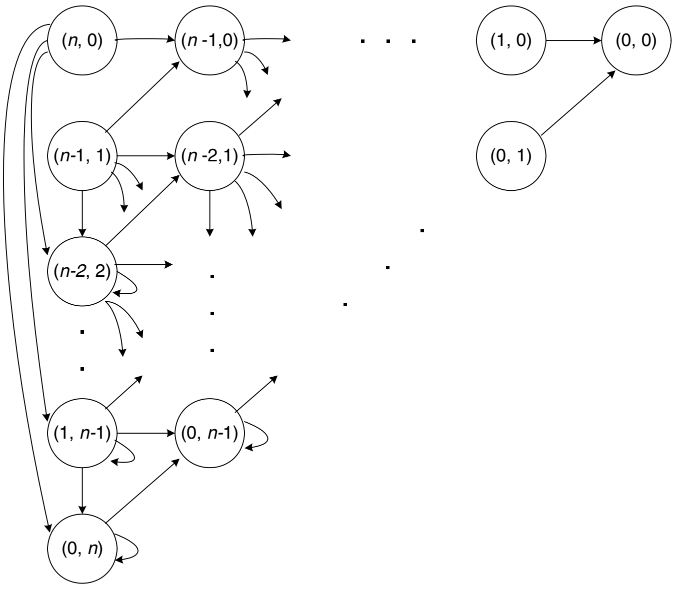


Fig. 3. The state transition diagram for (n_1, n_2) in a contention period.

for the contention window size. Define state (n_1, n_2), where n_1 and n_2 are the numbers of voice sessions with contention windows CW_1 and CW_2, respectively. Hence, in each CP, the initial state is $(n, 0)$. When a voice session contends successfully (i.e., it is the only one with the largest backoff timer), it will leave the contention. If there is a collision (i.e., there are at least two voice stations with the largest backoff timer among all the contending voice stations), the involved voice stations will double their contention window until the maximum contention window size (i.e., CW_2) is reached. After each successful transmission or collision, the state will evolve, remaining in the current state, or moving to the next one. The state transition is shown in Fig. 3, where the state $(0, 0)$ is the absorbing state when all voice sessions are served. There are totally $1+2+...+(n+1) = \frac{(n+1)(n+2)}{2}$ states. To understand the state transition diagram, we use state $(1, n-1)$ as an example. Its next state is $(0, n-1)$ if the voice session with CW_1 transmits successfully, $(1, n-2)$ if one voice session with CW_2 transmits successfully, $(0, n)$ if the session with CW_1 collides with one or more other sessions, or it remains in $(1, n-1)$ if two or more sessions with CW_2 collide. From the diagram we can also see that the probability of staying in state $(n-1, 1)$ is 0, as no other state enters it.

Let $T(n_1, n_2)$ denote the average time needed for transitions from state (n_1, n_2) to the absorbing state (0, 0). Obviously, we have $T(0, 0) = 0$, and $T(n, 0)$ is the average time to serve all the n contending voice sessions in a CP.

For a state (n_1, n_2), one or more transmissions from voice sessions with either CW_1 or CW_2 will lead to its next state. Denote the number of transmissions from voice sessions with CW_1 and CW_2 as $l_1(\leq n_1)$ and $l_2(\leq n_2)$, respectively. Denote the next state as $s_{(n_1,n_2;\ l_1,l_2)}$. Then we have

$$s_{(n_1,n_2;\ l_1,l_2)} = \begin{cases} (n_1-l_1, n_2-l_2) & \text{if } l_1+l_2 = 1 \\ (n_1-l_1, n_2+l_1) & \text{if } l_1+l_2 > 1 \end{cases} \quad (1)$$

where $l_1+l_2 = 1$ means a successful transmission. When $l_1 + l_2 > 1$, a collision happens, and the l_1 involved voice stations originally with CW_1 will be with CW_2 after the collision.

Denote the probability of the above transition as $p_{(n_1,n_2;\ l_1,l_2)}$, and the average time of the transition as $t_{(n_1,n_2;\ l_1,l_2)}$. If $l_1 \neq 0$, i.e., the successful transmission or collision happens when the largest backoff timer among all the voice stations takes a value from [0, CW_1], we have

$$p_{(n_1,n_2;\ l_1,l_2)} = \sum_{i=0}^{CW_1} \binom{n_1}{l_1} \left(\frac{1}{CW_1+1}\right)^{l_1} \left(\frac{i}{CW_1+1}\right)^{n_1-l_1} \cdot \binom{n_2}{l_2} \left(\frac{1}{CW_2+1}\right)^{l_2} \left(\frac{i}{CW_2+1}\right)^{n_2-l_2} \quad (2)$$

where the term in the summation means the probability that l_1 voice stations with CW_1 and l_2 voice stations with CW_2 choose a backoff timer value i, and other voice stations choose backoff timer values less than i.

With the condition of the above transition, the conditional probability that the largest backoff timer value in the successful transmission or collision is i can be given by

$$\frac{\binom{n_1}{l_1}\left(\frac{1}{CW_1+1}\right)^{l_1}\left(\frac{i}{CW_1+1}\right)^{n_1-l_1} \cdot \binom{n_2}{l_2}\left(\frac{1}{CW_2+1}\right)^{l_2}\left(\frac{i}{CW_2+1}\right)^{n_2-l_2}}{p_{(n_1,n_2;\ l_1,l_2)}}$$

and we have

$$t_{(n_1,n_2;\ l_1,l_2)} = \sum_{i=0}^{CW_1} \frac{1}{p_{(n_1,n_2;\ l_1,l_2)}} \cdot \binom{n_1}{l_1} \left(\frac{1}{CW_1+1}\right)^{l_1} \left(\frac{i}{CW_1+1}\right)^{n_1-l_1} \cdot \binom{n_2}{l_2} \left(\frac{1}{CW_2+1}\right)^{l_2} \left(\frac{i}{CW_2+1}\right)^{n_2-l_2} \cdot i \cdot \tau + \tau + T_x \quad (3)$$

where τ is the slot duration. On the right side of (3), the first term (i.e., the summation) represents the time used by the busy tone, the second term (i.e., τ) is the duration for busy tone detection after a node finishes its own busy tone, and the third term (i.e., T_x) is the collision or successful transmission time, including the AIFS[AC_voice], the packet transmission time, SIFS and ACK transmission time for a successful transmission (when $l_1 + l_2 = 1$), or ACK timeout for a collision (when $l_1 + l_2 > 1$).

If $l_1 = 0$, the transmission or collision can happen when the largest backoff timer among all the contending voice stations takes a value from [0, CW_2]. We have

$$p_{(n_1,n_2;\ 0,l_2)} = \sum_{i=0}^{CW_1} \left(\frac{i}{CW_1+1}\right)^{n_1} \cdot \binom{n_2}{l_2} \left(\frac{1}{CW_2+1}\right)^{l_2} \left(\frac{i}{CW_2+1}\right)^{n_2-l_2} + \sum_{i=CW_1+1}^{CW_2} \binom{n_2}{l_2} \left(\frac{1}{CW_2+1}\right)^{l_2} \left(\frac{i}{CW_2+1}\right)^{n_2-l_2}$$

$$t_{(n_1,n_2;\ 0,l_2)} = \left[\sum_{i=0}^{CW_1} \frac{\left(\frac{i}{CW_1+1}\right)^{n_1} \cdot \binom{n_2}{l_2}\left(\frac{1}{CW_2+1}\right)^{l_2}\left(\frac{i}{CW_2+1}\right)^{n_2-l_2}}{p_{(n_1,n_2;\ 0,l_2)}} \cdot i \cdot \tau + \sum_{i=CW_1+1}^{CW_2} \frac{\binom{n_2}{l_2}\left(\frac{1}{CW_2+1}\right)^{l_2}\left(\frac{i}{CW_2+1}\right)^{n_2-l_2}}{p_{(n_1,n_2;\ 0,l_2)}} \cdot i \cdot \tau\right] + \tau + T_x. \quad (4)$$

$$T(n_1,n_2) = \sum_{0\le l_1\le n_1,\ 0\le l_2\le n_2,\ l_1+l_2>0} p_{(n_1,n_2;\ l_1,l_2)}[T(s_{(n_1,n_2;\ l_1,l_2)}) + t_{(n_1,n_2;\ l_1,l_2)}] \tag{5}$$

Hence, consider all possible transitions from state (n_1, n_2) where $n_1 + n_2 > 0$, we have (5) at the top of this page. From (5) and $T(0, 0) = 0$, we can compute the values of $T(n_1, n_2)$.

In addition, to implement this analytical work in a practical system, a lookup table can be generated in advance to reduce the computation complexity in system configuration.

B. Packet Loss Rate Bound in a CFP

It is critical to determine how many voice sessions can be supported in the CFP with QoS guarantee. The delay requirement of voice service can be guaranteed by the controlled access in the CFP. Hence, our focus is the packet loss rate in the CFP.

Consider N voice sessions at a service interval. Each voice session has the independent on and off periods exponentially distributed with mean values $1/\alpha$ and $1/\beta$, respectively. At a time instant, a voice station is at on state with probability $\frac{\beta}{\alpha+\beta}$, and at off state with probability $\frac{\alpha}{\alpha+\beta}$. When a voice station is at on state, the probability that a transition to off state happens after duration t is given by $\exp(-\alpha \cdot t)$. When a voice station is at off state, the probability that a transition to on state happens after duration t is given by $\exp(-\beta \cdot t)$. The maximum number of downlink (or uplink) voice packets generated in a service interval from a voice session is $M = T_{SI}/t_a$. For each voice session, let $P_d(i)$ and $P_u(i)$ $(0 \le i \le M)$ denote the probability of generating i downlink and uplink voice packets respectively for transmission in a CFP. We have

$$P_d(i) = \begin{cases} \frac{\alpha}{\alpha+\beta}[\exp(-\beta\cdot(T_{SI} - i\cdot t_a)) \\ \quad - \exp(-\beta\cdot(T_{SI} - (i-1)\cdot t_a))] \\ \quad + \frac{\beta}{\alpha+\beta}[\exp(-\alpha(i-1)t_a) - \exp(-\alpha\cdot i\cdot t_a)] \\ \qquad\qquad 1 \le i \le M-1 \\ \frac{\alpha}{\alpha+\beta}[1-\exp(-\beta\cdot t_a)] + \frac{\beta}{\alpha+\beta}\exp(-\alpha(T_{SI}-t_a)) \\ \qquad\qquad i = M \\ 1 - \sum_{i=1}^{M} P_d(i) \qquad i = 0 \end{cases} \tag{6}$$

and

$$P_u(i) = \begin{cases} \frac{\beta}{\alpha+\beta}[\exp(-\alpha(i-1)t_a) \\ \quad - \exp(-\alpha\cdot i\cdot t_a)] & 1 \le i \le M-1 \\ \frac{\beta}{\alpha+\beta}\exp(-\alpha(T_{SI}-t_a)) & i = M \\ 1 - \sum_{j=1}^{M} P_u(j) & i = 0. \end{cases} \tag{7}$$

We use the example of $1 \le i \le M-1$ to explain the above equations. For the downlink, a voice session will generate i packets when it is originally (at the beginning of the service interval) off with probability $\frac{\alpha}{\alpha+\beta}$ and transits to the on state within $[T_{SI} - i\cdot t_a, T_{SI} - (i-1)\cdot t_a)$ with probability $[\exp(-\beta\cdot(T_{SI}-i\cdot t_a)) - \exp(-\beta\cdot(T_{SI}-(i-1)\cdot t_a))]$, or when it is originally on with probability $\frac{\beta}{\alpha+\beta}$ and transits to the off state within $((i-1)t_a, i\cdot t_a]$ with probability $[\exp(-\alpha(i-1)t_a) - \exp(-\alpha\cdot i\cdot t_a)]$. The uplink case is different from the downlink case, as the first several packets in each talk spurt (when an off to on transition happens within T_{SI}) in the uplink are transmitted in the CP.

Next we estimate the number (X) of voice packets that can be supported in each CFP. We call the one-way (i.e., uplink or downlink) packets of a voice session ready for transmission in the CFP *a burst* (which will be transmitted by a single MAC frame). For a burst in a CFP, the probability that it is an uplink transmission with size $i(1 \le i \le M)$ is

$$\frac{P_u(i)}{\sum_{i=1}^{M} P_d(i) + \sum_{i=1}^{M} P_u(i)}$$

and the probability that it is a downlink transmission with size $i(1 \le i \le M)$ is

$$\frac{P_d(i)}{\sum_{i=1}^{M} P_d(i) + \sum_{i=1}^{M} P_u(i)}.$$

Then the probability that it is in uplink transmission (thus requiring a CF-Poll) is

$$\frac{\sum_{i=1}^{M} P_u(i)}{\sum_{i=1}^{M} P_d(i) + \sum_{i=1}^{M} P_u(i)}.$$

The average burst size is given by

$$B = \sum_{i=1}^{M} \frac{P_u(i)}{\sum_{j=1}^{M} P_d(j) + \sum_{j=1}^{M} P_u(j)} \cdot i + \sum_{i=1}^{M} \frac{P_d(i)}{\sum_{j=1}^{M} P_d(j) + \sum_{j=1}^{M} P_u(j)} \cdot i. \tag{8}$$

The average number of bursts is X/B.

We have

$$T_{CFP} = (X/B)\cdot(T_o + \frac{L\cdot B}{R}) + (X/B)\frac{\sum_{i=1}^{M} P_u(i)}{\sum_{i=1}^{M} P_d(i) + \sum_{i=1}^{M} P_u(i)} \cdot T_{poll} \tag{9}$$

where T_{CFP} is the duration of the CFP, T_o is the overhead due to IFS, PHY preamble, and MAC overhead, L is the payload size of a voice packet, R is the transmission rate of voice payload, and T_{poll} is the polling overhead. Then

$$X = B \cdot \frac{T_{CFP}}{T_o + \frac{L\cdot B}{R} + \frac{\sum_{i=1}^{M} P_u(i)}{\sum_{i=1}^{M} P_d(i) + \sum_{i=1}^{M} P_u(i)} \cdot T_{poll}}. \tag{10}$$

Let X_i denote the total number of up- and downlink voice packets from the voice session i ready for transmission in the CFP of a service interval, and $Y = \sum_{i=1}^{N} X_i$, where N is the total number of voice sessions. The expectation E$[X_i]$ and variance Var$[X_i]$ of X_i can be determined based on the on/off model. If the packet loss rate in the CFP is required to be bounded by P_L, the following inequality should hold:

$$\frac{\sum_{Y>X}(Y-X)P(Y)}{E[Y]} \le P_L \tag{11}$$

where $P(Y)$ is the probability mass function of Y. According to the central limit theorem, the random variable Y can be approximated as a Gaussian random variable with mean

$N \cdot \mathrm{E}[X_i]$ and variance $N \cdot \mathrm{Var}[X_i]$ when N is large. The maximum N satisfying the above inequality (11) is the maximum voice session number supported by the CFP.

C. Session Admission Control for Voice

In order to guarantee QoS of voice traffic, it is critical to have an appropriate session admission control algorithm. The AP is responsible for admitting or rejecting a new voice session based on the available resources to ensure that the QoS requirements (such as delay and packet loss rate) of all the admitted voice sessions are satisfied. For session admission control, it is essential to have the capacity region. IEEE 802.11e has given a reference design. When there are k existing voice sessions in a BSS, a new voice session indexed by $k+1$ can be admitted if the following inequality holds:

$$\frac{TXOP_{k+1}}{T_{SI}} + \sum_{i=1}^{k} \frac{TXOP_i}{T_{SI}} \leq 1 - \rho_{CP} \tag{12}$$

where ρ_{CP} is the minimum percentage of time used for the CP during each beacon interval and $TXOP_i$ the minimum time that needs to be allocated for session i to ensure its QoS requirements. The value of $TXOP_i$ depends on the voice packet size and the packet arrival rate. This algorithm is suitable only for constant-rate voice traffic without statistical multiplexing. Based on the algorithm, variable-rate voice traffic (represented by the `on/off` model) requires much more resources than what is actually needed. Here, we propose another algorithm to determine the capacity region which takes into account statistical multiplexing and at the same time guarantees the delay and packet loss rate requirements of voice traffic.

For a WLAN supporting voice/data traffic, the QoS requirement of the low-priority data traffic should also be guaranteed. In our system, data traffic is guaranteed the average service time in each service interval, i.e., an average fraction ϕ of time in a service interval is used by data traffic in the long run. Hence, we need to determine how many voice sessions can be admitted with average service time $(1-\phi)T_{SI}$ (used in both CFP and CP) in each service interval, and with the required packet loss rate guaranteed. Let $\overline{T_{CP}^{v}}$ denote the average time in a CP used to serve contending voice sessions. Thus,

$$T_{CFP} + \overline{T_{CP}^{v}} \leq (1-\phi)T_{SI}. \tag{13}$$

In Section IV-A, we derive the average time required in a CP to serve a fixed number, n, of voice sessions contending in the CP. However, with N voice sessions in service, the number of contending voice sessions in a CP varies (due to the voice `on/off` nature), so does the required service time in the CP. The average service time for contending voice sessions in a CP is given by

$$\overline{T_{CP}^{v}}(N) = \sum_{n=1}^{N} \binom{N}{n} (P_c)^n \cdot (1-P_c)^{N-n} \cdot T(n,0) \tag{14}$$

where $T(n,0)$ is the average time to serve n contending voice sessions in a CP (as defined in Section IV-A), and P_c is the probability that a voice station contends for the channel in a CP, given by

$$P_c = \frac{\alpha}{\alpha+\beta}[1 - \exp(-\beta \cdot T_{SI})]. \tag{15}$$

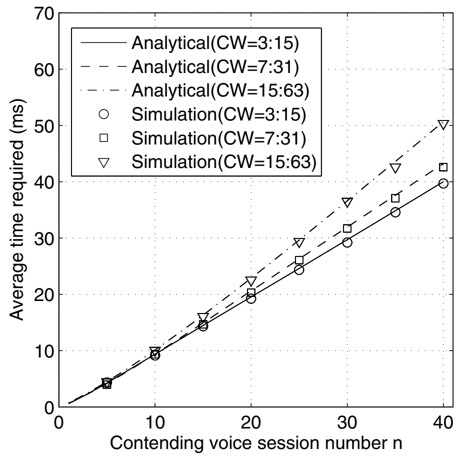


Fig. 4. Average time required to serve contending voice sessions in a CP in our scheme.

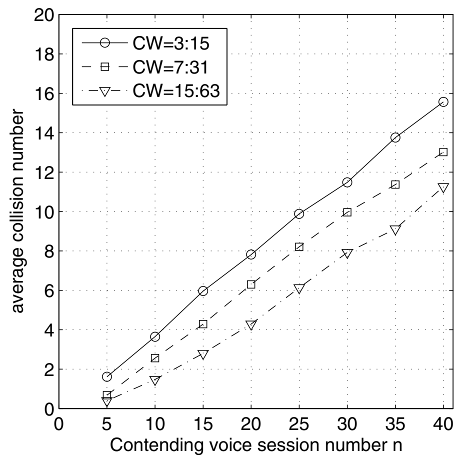


Fig. 5. Average voice collision number (via simulations) in a CP in our scheme.

In this work, the duration of the CP in a service interval, $T_{CP} = T_{SI} - T_{CFP}$, is larger than $\overline{T_{CP}^{v}}(N)$, as the difference of them is the average service time for data traffic in a CP. The performance of contending voice sessions in a CP can be evaluated by the outage probability that T_{CP} is not sufficient to serve all the contending voice sessions

$$\xi(N) = \sum_{n=1}^{N} \binom{N}{n} (P_c)^n \cdot (1-P_c)^{N-n} \cdot I\{T(n,0) > T_{CP}\} \tag{16}$$

where $I\{\cdot\}$ is an indicator function.

From the analysis in Section IV-B, if N voice sessions are admitted, we can determine the minimum value of T_{CFP}, denoted by $T_{CFP}^{m}(N)$, in order to guarantee the voice packet loss rate bound. From the constraint (13), the capacity region for voice is the maximum integer N (denoted by N^*) satisfying

$$T_{CFP}^{m}(N) + \overline{T_{CP}^{v}}(N) \leq (1-\phi)T_{SI} \tag{17}$$

and the service interval should be configured with a CFP with duration $T_{CFP}^{m}(N^*)$ and a CP with duration $T_{SI} - T_{CFP}^{m}(N^*)$.

TABLE II
SIMULATION PARAMETERS

Parameter	Value
Slot time τ	20 μs
T_{SI}	100 ms
SIFS	10 μs
AIFS[AC_voice]	40 μs
AIFS[AC_data]	60 μs
PHY preamble	192 μs
MAC header	34 bytes
ACK	14 bytes
CF-Poll	36 bytes
R	11 Mbps
Basic rate	2 Mbps
$1/\alpha$	352 ms
$1/\beta$	650 ms
t_a	20 ms
L	33 bytes
Data packet payload	1000 bytes
P_L	1%

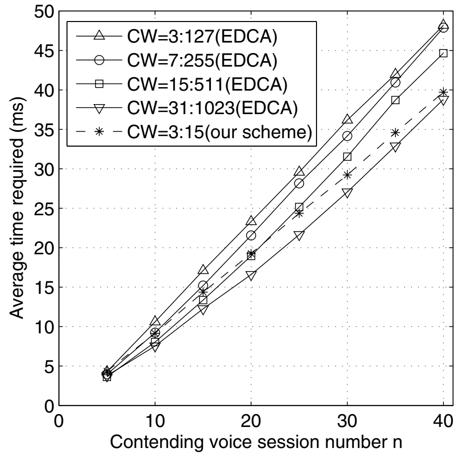


Fig. 6. Average time required to serve voice sessions in the CP in EDCA compared with our scheme.

V. Numerical Results and Discussion

To validate the analysis and evaluate the performance of our proposed scheme, computer simulations are carried out using Matlab. The simulation for each run consists of 1000 service intervals. We choose the GSM 6.10 codec as the voice source as an example. The voice payload size is 33 bytes and the packet inter-arrival period is 20 ms. Compressed RTP/UDP/IP headers with size 4 bytes are used in all the simulations. Other simulation parameter values are listed in Table II. We first vary the contending voice session number (i.e., n) in a CP (where each voice session has one MAC frame to send), and analyze and/or simulate the time to serve all the MAC frames (i.e., the time to serve the contending voice sessions in a CP). Then we evaluate how the packet loss rate in a CFP changes with the number of voice sessions N. In the evaluation, the first several packets of an uplink talk spurt are not transmitted in the CFP (but are transmitted in the CP by contention). Finally we evaluate the capacity of the whole system, and compare it with that of IEEE 802.11e. We obtain the portion of time required to serve different number of admitted voice sessions, and obtain the system capacity.

A. Time to Serve Contending Voice Sessions in a CP

For uplink voice transmission in our scheme, the first several packets in each talk spurt are transmitted in the CP. With the system parameters, the probability of a voice session contending in a CP is around 9% according to the analysis. Hence, if the total voice session number is 200, there are on average 18 voice sessions contending in each CP. Fig. 4 and Fig. 5 show the average time required to serve contending voice sessions (i.e., $T(n, 0)$) and average voice collision number in a CP versus the contending voice session number n with different settings for initial and maximum contention windows ($CW_{\min}$ and $CW_{\max}$), respectively. It is clear that our analysis matches well with the simulations. Contention window settings are critical for contention-based channel access. In our scheme, when the voice sessions have

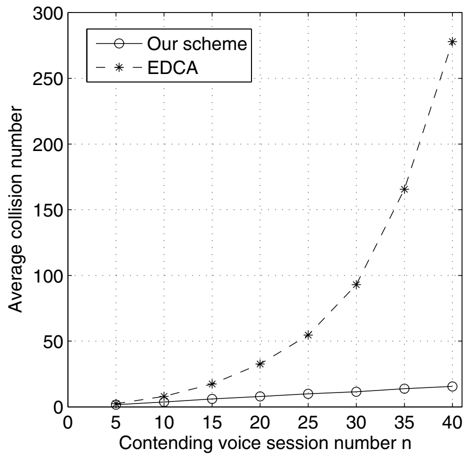


Fig. 7. Average collision number in a CP when our scheme or EDCA is applied with $CW_{\min} = 3$ and $CW_{\max} = 15$.

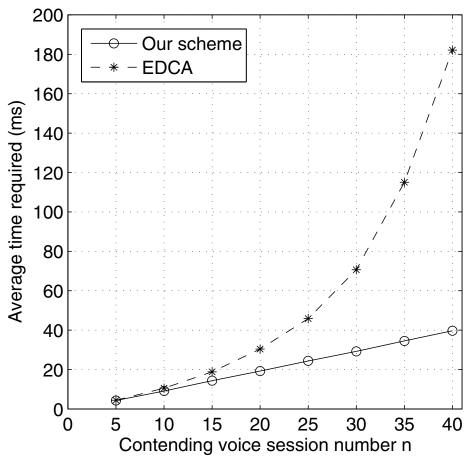


Fig. 8. Average time required to serve contending voice sessions in a CP when our scheme or EDCA is applied with $CW_{\min} = 3$ and $CW_{\max} = 15$.

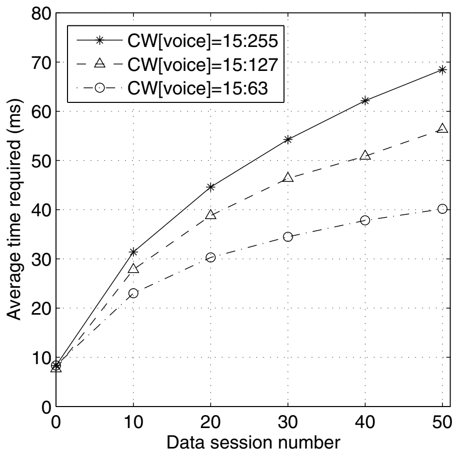


Fig. 9. Average time required so that all contending voice sessions can be served in a CP when EDCA is used to support 10 voice sessions and variable number of long-lived data sessions with CW[data]=31 : 1023.

smaller $CW_{\min}$ and $CW_{\max}$, the time to transmit busy tone is smaller, at the cost of more collisions. Via the analysis and simulations, we find out that $CW_{\min} = 3$ and $CW_{\max} = 15$ can lead to the minimal average time to serve all the voice sessions. Via simulations, we also compare this best case with the cases if EDCA of IEEE 802.11e is applied in the CP, and demonstrate the results in Fig. 6. It can be seen that there is no much difference between our scheme and EDCA. Only EDCA with $CW_{\min} = 31$ and $CW_{\max} = 1023$ has a non-trivial gain over our proposed scheme. However, to obtain priority in EDCA, voice AC is very likely to have a smaller $CW_{\min}$ (< 31) and $CW_{\max}$ (< 1023). Although a voice station with the largest backoff timer instead of the smallest one (as in EDCA) gets the channel in our scheme, voice performance is not degraded much. The reason is that our proposed scheme can use very small $CW_{\min}$ and $CW_{\max}$, but EDCA cannot. If our proposed scheme and EDCA use the same $CW_{\min}$ and $CW_{\max}$, the backoff waiting time in our scheme is larger. However, our scheme has a smaller collision probability. If multiple nodes choose the same backoff timer, a collision will occur in EDCA, but a collision will happen in our scheme only when the multiple nodes are with the largest backoff timer (among all the nodes). Fig. 7 and Fig. 8 show the average collision number and required time to serve contending voice sessions, respectively, in EDCA and our scheme with $CW_{\min} = 3$ and $CW_{\max} = 15$ via simulations. It can be seen that, as the contending voice session number increases, the collision number increases rapidly in EDCA, but relatively slowly in our scheme. Hence, the time required to serve contending voice sessions in our proposed scheme is much smaller than that in EDCA when the contending session number is large in the example.

In addition, in the above comparison, EDCA is applied with no contending data sessions. As discussed in Section II-B, when data sessions are added, the voice performance of EDCA will be degraded, but our scheme will not be affected because it can guarantee deterministic priority to voice sessions over data sessions. Fig. 9 shows the effect of data traffic on the average time required in a CP so that all contending voice

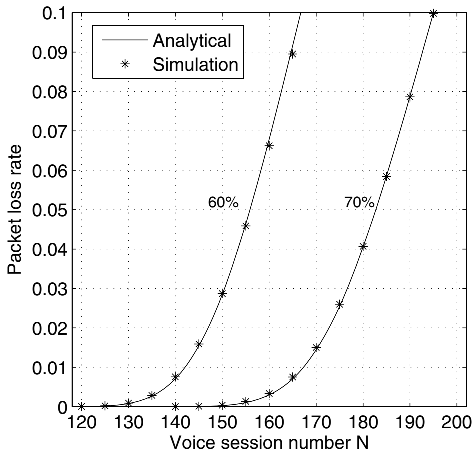


Fig. 10. Packet loss rate in CFP for $T_{CFP} = 60\% T_{SI}$ and $T_{CFP} = 70\% T_{SI}$ in our scheme.

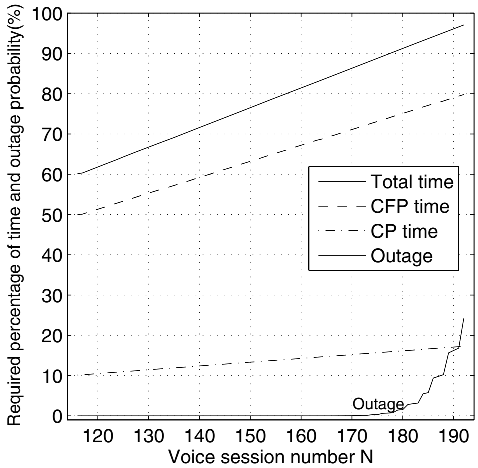


Fig. 11. The percentage of time (in CP, CFP, and totally) in a service interval needed to serve the voice sessions with QoS guarantee and the outage probability that not all the contending voice sessions can be served in a CP.

sessions can be served in EDCA. Long-lived data sessions use the initial and maximum contention window pair $(31, 1023)$, while voice sessions choose initial contention window size 15 and maximum contention window size 63, 127, or 255. We can see that when the number of data sessions increases, the average time required increases accordingly, and the negative effect is more significant if voice sessions choose a larger initial and maximum contention window pair ($(15, 255)$ in the example).

B. Packet Loss Rate in the CFP

Fig. 10 shows the analytical results of the packet loss rate versus voice session number N with T_{CFP} equal to 60% and 70% of T_{SI} in our scheme. Simulations are also carried out for selected values of N. It can be seen that our analysis matches well with the simulations. From Fig. 10, when the voice session number is equal to or less than 141 when $T_{CFP} = 60\% T_{SI}$ or 167 when $T_{CFP} = 70\% T_{SI}$, the packet loss rate in the CFP is bounded by 1%.

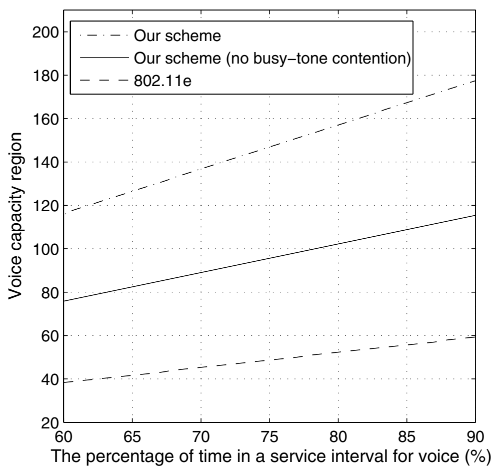


Fig. 12. The analyzed voice traffic capacity region of our proposed scheme, our scheme without busy-tone contention mechanism, and IEEE 802.11e.

C. Capacity Region of Voice

To determine the capacity region of voice in our scheme, we vary the number N of voice sessions in the system, and calculate the average time in the CP $\overline{T_{CP}^{v}}(N)$ and the duration of the CFP $T_{CFP}^{m}(N)$ in order to guarantee that voice packet loss rate in the CFP is bounded by 1%. We further obtain the total average time in a service interval needed to serve the N voice sessions with QoS guarantee. The analytical results are shown in Fig. 11, which also gives the outage probability that the CP with duration $[T_{SI} - T_{CFP}^{m}(N)]$ cannot serve all contending voice sessions. It is shown that, when data traffic requires average 30% service time (thus 70% time for voice) in a service interval, we should configure $T_{CFP} \approx 57\% T_{SI}$ and $T_{CP} \approx 43\% T_{SI}$ with a maximum admitted voice session number of 136. The outage probability is negligible ($< 1\%$) if the total average time for voice is less than 90%.

We further obtain the voice capacity region (i.e., the maximum number of voice sessions that can be admitted) when the percentage of time (in both CFP and CP) used by voice in each service interval varies from 60% to 90%, and get the analytical results as shown in Fig. 12. The analyzed voice capacity region of the IEEE 802.11e polling scheme with the same percentage of time for voice is also included in Fig. 12. For a comparison, Fig. 12 also shows the voice capacity region when our scheme is applied without the busy-tone contention mechanism (i.e., with only the overhead reduction mechanism). For uplink transmissions, all voice stations are polled, and if a polled voice station has no packets to transmit, it will respond with a NULL frame. From Fig. 12, it can be seen that our proposed overhead reduction and busy-tone contention mechanisms both can significantly improve system capacity as compared with IEEE 802.11e.

VI. Conclusion

To support real-time voice traffic as well as data traffic over WLANs, the controlled channel access is preferred to the contention-based access for voice. In this paper, we propose solutions to enhance QoS provisioning capability of IEEE 802.11e to guarantee the delay requirement of voice and at the same time achieve bandwidth efficiency and a certain level of QoS assurance for data. Voice statistical multiplexing gain is exploited effectively, and the system overhead is reduced significantly. We provide an analytical model for the contention period and contention-free period, which is validated by extensive simulations. A session admission control algorithm is presented to admit voice sessions into the system. Our solutions are shown to significantly improve the capacity of IEEE 802.11e WLANs supporting voice and data. This research should provide helpful insights to the development and deployment of VoIP technologies over WLANs (which were originally designed for data services).

Acknowledgements

The authors would like to thank the anonymous reviewers for their constructive comments which improve the presentation of this paper.

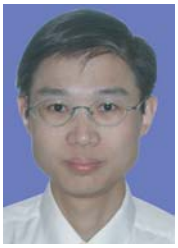
References

[1] W. Wang, S. C. Liew, and V. O. K. Li, "Solutions to performance problems in VoIP over a 802.11 wireless LAN," *IEEE Trans. Veh. Technol.*, vol. 54, no. 1, pp. 366–384, Jan. 2005.
[2] D. P. Hole and F. A. Tobagi, "Capacity of an IEEE 802.11b wireless LAN supporting VoIP," in *Proc. IEEE ICC'04*, pp. 196-201, June 2004.
[3] S. Garg and M. Kappes, "An experimental study of throughput for UDP and VoIP traffic in IEEE 802.11b networks," in *Proc. IEEE WCNC'03*, pp. 1748-1753, Mar. 2003.
[4] Q. Ni, L. Romdhani, and T. Turletti, "A survey of QoS enhancements for IEEE 802.11 wireless LAN," *Wireless Commun. Mobile Comput.*, vol. 4, no. 5, pp. 547–566, Aug. 2004,
[5] X. Yang and N. H. Vaidya, "Priority scheduling in wireless ad hoc networks," in *Proc. ACM MOBIHOC' 02*, pp. 71–79.
[6] J. W. Robinson and T. S. Randhawa, "Saturation throughput analysis of IEEE 802.11e enhanced distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 5, pp. 917–928, June 2004.
[7] D. Chen, S. Garg, M. Kappes, and K. Trivedi, "Supporting VBR VoIP traffic in IEEE 802.11 WLAN in PCF mode," Avaya Laboratories, Basking Ridge, NJ, Tech. Rep. ALR-2002-026, 2002.
[8] M. Veeraraghavan, N. Cocker, and T. Moors, "Support of voice services in IEEE 802.11 wireless LANs," in *Proc. IEEE INFOCOM'01*, pp. 488-497.
[9] E. Ziouva and T. Antonakopoulos, "A dynamically adaptable polling scheme for voice support in IEEE802.11 networks," *Computer Commun.*, vol. 26, no. 2, pp. 129–142, Feb. 2003.
[10] Y. Xiao, H. Li, and S. Choi, "Protection and guarantee for voice and video traffic in IEEE 802.11e wireless LANs," in *Proc. IEEE INFOCOM'04*, pp. 2152–2162.
[11] H. Kim and J. C. Hou, "Improving protocol capacity for UDP/TCP traffic with model-based frame scheduling in IEEE 802.11-operated WLANs," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 10, pp. 1987–2003, Dec. 2004
[12] R. O. Bladwin, N. J. Davis IV, S. F. Midkiff, and R. A. Raines, "Packetized voice transmission using RT-MAC, a wireless real-time medium access control protocol," *Mobile Comput. Commun. Review*, vol. 5, no. 3, pp. 11–25, July 2001.
[13] W. F. Fan, D. H. K. Tsang, and B. Bensaou, "Admission control for variable bit rate traffic using variable service interval in IEEE 802.11e WLANs," in *Proc. IEEE ICCCN'04*, pp. 447–453.
[14] Y.-L. Kuo, C.-H. Lu, E. H.-K. Wu, and G.-H. Chen, "An admission control strategy for differentiated services in IEEE 802.11," in *Proc. IEEE GLOBECOM'03*, pp. 707–712.
[15] S. Garg and M. Kappes, "Admission control for VoIP traffic in IEEE 802.11 networks," in *Proc. IEEE GLOBECOM'03*, pp. 3514–3518.
[16] D. Pong and T. Moors, "Call admission control for IEEE 802.11 contention access mechanism," in *Proc. IEEE GLOBECOM'03*, pp. 174–178.
[17] J. L. Sobrinho and A. S. Krishnakumar, "Quality-of-service in ad hoc carrier sense multiple access wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 8, pp. 1353–1368, Aug. 1999.

[18] S. Casner and V. Jacobson, "Compressing IP/UDP/RTP Headers for Low-Speed Serial Links," IETF RFC 2508, Feb. 1999.

[19] L.-A. Larzon, H. Hannu, L.-E. Jonsson, and K. Svanbro, "Efficient transport of voice over IP over cellular links," in *Proc. IEEE GLOBECOM'00*, pp. 1669–1676.

Ping Wang received the B.E. and M.E. degrees in 1994 and 1997, respectively, both in electrical engineering, from Huazhong University of Science and Technology, Wuhan, China. She is currently working toward a Ph.D. degree at the Department of Electrical and Computer Engineering, University of Waterloo, Canada. Her current research interests include QoS provisioning and resource allocation in multimedia wireless communications.



Hai Jiang (M'07) received the B.S. degree in 1995 and the M.S. degree in 1998, both in electronics engineering, from Peking University, Beijing, China, and the Ph.D. degree in 2006 in electrical engineering from the University of Waterloo, Canada. He is currently a Postdoctoral Fellow at the Department of Electrical Engineering, Princeton University. His research interests include radio resource management, cellular/WLAN interworking, and cross-layer design for wireless multimedia communications.

Weihua Zhuang (M'93-SM'01) received the B.Sc. and M.Sc. degrees from Dalian Maritime University, Liaoning, China, and the Ph.D. degree from the University of New Brunswick, Fredericton, NB, Canada, all in electrical engineering. Since October 1993, she has been with the Department of Electrical and Computer Engineering, University of Waterloo, ON, Canada, where she is a full professor. She is a co-author of the textbook *Wireless Communications and Networking* (Prentice Hall, 2003). Her current research interests include multimedia wireless communications, wireless networks, and radio positioning. Dr. Zhuang is a licensed Professional Engineer in the Province of Ontario, Canada. She received the Premier's Research Excellence Award (PREA) in 2001 from the Ontario Government for demonstrated excellence of scientific and academic contributions. She is the Editor-in-Chief of *IEEE Transactions on Vehicular Technology*, and an Editor of *IEEE Transactions on Wireless Communications*, *EURASIP Journal on Wireless Communications and Networking*, and *International Journal of Sensor Networks*.