Can I add a VoIP call?

Sachin Garg
Avaya Labs
Basking Ridge, NJ 07920
Email: sgarg@avaya.com

Martin Kappes
Avaya Labs
Basking Ridge, NJ 07920
Email: mkappes@avaya.com

***Abstract*— In this paper, we study the inherent limitations of the 802.11 (a/b) distributed coordination function (DCF) in supporting VoIP calls over a wireless LAN. Specifically, we evaluate the upper bound on the number of simultaneous VoIP calls that can be placed in a single cell of an 802.11(a/b) network. Making one additional VoIP call in that cell would degrade the quality of all VoIP calls. The upper bound is calculated as a function of the choice of VoIP codec and the length of the audio payload. As an example, when a G711 codec with 20 millisecond audio payload is used, an 802.11b cell can support only 3 to 12 simultaneous VoIP calls. The actual number depends on the effective transmission rate of the wireless station, which for 802.11b can be 1Mbps, 2Mbps, 5.5Mbps and 11Mbps. We also study the effect of spatial distribution of the wireless stations on the upper bound which is the dominant factor in determining the effective transmission rate of a station.**

I. Introduction

The IEEE 802.11 wireless LANs [12] have been commoditized for data applications in enterprise networks. Meanwhile, VoIP deployment in an enterprise mostly involves VoIP terminals connected to its wireline IP network. However, as VoIP gains more traction, wireless VoIP terminals will also begin to be deployed to leverage the pervasive 802.11 networks for caller mobility.

Ensuring the Quality-of-Service for VoIP in wireless LANs is a big concern as the performance characteristics of their Physical and MAC layers is much worse than their wireline counterparts. In particular, lower peak transmission rate, lossy medium, interference problems etc. are some drawbacks of the PHY layer. At the MAC layer, 802.11 offers two choices. The first, called Distributed Coordination Function (DCF) belongs to the CSMA/CA family of protocols, where each station determines the access to the channel on its own without the involvement of any central coordinator. The second is called the Point Coordination Function (PCF) in which the access point (henceforth referred to as AP) determines which connected station gets to transmit at any time. While the PCF was designed to support real-time services, its implementation in an access-point and/or client cards is optional. As a result, the majority of 802.11(b) networks currently deployed do not support PCF.

As wireless VoIP takes hold, especially in public hot-spots such as airports, hotels etc., the question of how suitable DCF is for supporting VoIP traffic becomes important. That is the focus of this paper. In particular, we analyze the number of simultaneous VoIP calls a single AP running DCF can support. A naive calculation might indicate that given the 11Mbps peak rate and 128Kbps needed for a full duplex VoIP call, approximately 85 simultaneous calls can supported. The real motivation for the analysis comes from a simple experiment, the goal of which was to determine that number. We briefly describe the experiment, setup and its *surprising* outcome before moving on to the analysis.

In the experiment, multiple Wireless PCs running Windows 2000, were associated with the same 802.11b AP, which was connected to a 100Mbps Ethernet. The setup was used to make full-duplex VoIP calls between a wireless PC and a wired PC using IP phones. For each call, we used the ITU G711 a-Law codec where frames are sent out every 10 milliseconds. Each call results in two RTP streams, from wired to wireless and vice-versa. We tested the number of VoIP connection with acceptable voice quality by successively establishing new calls in addition to the ongoing calls. The quality of the connections was monitored through measurements of loss, jitter and round-trip time by a commercially available tool.

For the first five calls, the quality of all the calls was fine. Loss (0%), round-trip time (around 5 ms) and jitter (around 7 ms) were all in acceptable ranges for a good quality VoIP call. When the sixth call was placed, except for a sporadic increase in the round-trip time for some of the connections the quality of all six simultaneous connections was still acceptable. As soon as the seventh call was placed, all seven “wired to wireless” streams started suffering approximately 16% loss and the call quality became unacceptable for all calls in this direction. All “wireless to wired” streams still exhibited acceptable call quality. In short, the outcome of the experiement indicated that given the codec setting, only six calls can be placed on a single 802.11b AP.

The rest of the paper provides an explanation of the experimental observation and is organized as follows. Section II consists of a brief background on DCF for the sake of self containment and also gives a short summary of the related work. In Section III, we develop the simple analytical model to determine the upper bound on the number of simultaneous VoIP connections over DCF and in Section III-A, the model is used to tabulate and explain the results for an 802.11b AP. In Section III-B, we evaluate the effect of spatial distribution of the stations within the cell on the upper bound. The base model also qualifies to evaluate the upper bound for 802.11a [13] as only some parameter values change in 802.11a while the underlying DCF remains unchanged. These results are tabulated and explained in Section III-C. The paper concludes in Section IV.

II. Related Work and Background

With respect to 802.11 networks, prior performance studies, such as [3], have mainly focussed on analyzing the behavior of the MAC protocol itself. Stochastic Petri Nets are used in [7] to model the behavior of DCF and then performance measures such as effective channel throughput are derived. Bianchi, in [2], used a Markov process to model DCF and evaluated the channel throughput, and frame loss as a function of the number of wireless stations. In all of the above, the objective has been to study the performance of DCF itself while being agnostic to any protocols running over the DCF MAC. Hence the assumptions made on the traffic load do not correspond to the load generated specifically by VoIP traffic.

Prior work that pertains to studying VoIP over 802.11 has focussed mainly on PCF as the protocol of choice, for instance in [4]. Veeraraghavan et. al., in [10], present an analysis of the delay and loss characteristics for voice traffic over PCF as a function of the inter-poll period. A simulation approach was used in [11] to study and analyze variations in the polling schemes for PCF.

Enhancements to the MAC protocols that are neither DCF nor PCF have also been proposed [9] so that real-time QoS guarantees can be provided to the VoIP calls. No one, to the best of our knowledge, has quantified the the maximum number of simultaneous VoIP calls that can be supported on a single AP running the basic, standards compliant DCF. The upper bound is in the sense that adding one more VoIP call will disrupt the quality of all ongoing calls. Further, as the purpose in this paper is to study the inherent limitations of DCF vis-a-vis VoIP, it is assumed that no data traffic is present. Presence of any data traffic will only reduce the number of simultaneous calls.

Before the analysis itself, we give a brief introduction to DCF as follows. As collisions in the wireless medium cannot be detected, the MAC protocol is designed to prevent collisions from occurring and required sensing of the wireless medium. All unicast frames are acknowledged by the receiving station within a certain duration of receiving the frame. This duration is called the short inter frame space (SIFS). A node may transmit a frame if it senses the medium idle for a certain duration of time called the DCF inter frame space (DIFS). Since DIFS is longer than SIFS, a correctly received frame is always acknowledged before the channel is used again.

If a node wants to start transmitting while the medium is busy or if it wants to transmit another frame after just finishing a transmission, it also waits for the medium to be idle for the DIFS period. Then, the node does not begin to transmit immediately but enters a contention phase for the medium. Contention is done by choosing an integer random backoff between 0 and a parameter CW (CW stands for contention window size) which is initialy set to a value CWmin. The probability distribution among these values is uniform. The random backoff determines the number of time slots the client defers its transmission in addition to the DIFS time.

If the medium is sensed idle in such a "slot", the backoff timer is decreased by one. If the random backoff has decreased to 0, the node starts transmitting. If another node starts transmitting before this happens, the node continues to count down the backoff timer after the medium has been sensed idle for the DIFS period. Thus, if multiple clients want to transmit a frame, the one with the lowest random backoff time will win the contention for the medium. However, if more than one node happens to choose the same backoff time, they will start to transmit at the same time and a collision will occur. The clients assume that the frame was lost if an acknowledgment is not received within SIFS. In case of an unsuccessful transmission, the CW value doubles until a CWmax value is reached. The CW parameter is reset to the CWmin after each successful (i.e. acknowledged) transmission.

The IEEE 802.11/802.11b standard defines SIFS to be 10 μs. A slot time is 20 μs and the value of DIFS is 50 μs. The size of an acknowledgment frame is 14 bytes which take about 10 μs to transmit at 11 Mbps. However, each transmitted frame also needs some physical layer overhead (PLCP header of 48 μs and a preamble of 144 μs) which is about 192 μs. Thus, the total time to transmit an acknowledgment is 203 μs. CWmin is defined to be 31.

III. Analysis for Number of VoIP connections

With the basic understanding of DCF, we move on to calculate the maximum number of VoIP clients a single AP can support. The analysis is based on the assumption that one end-point of each VoIP call is a wireless client, while the other end-point is on the wired network. This number depends on the maximum throughput the channel can achieve, which is a function of the packet size. Other factors effecting the channel throughput include the byte overheads of RTP, UDP, IP, MAC and physical layers. Further, the channel access mechanism (CSMA/CA) imposes an overhead due to the back-off procedures between successive packet transmissions from the same station.

Assume the following terminology. Let P be the size of voice payload. For G711 a-Law codec, this payload is 80 bytes for 10 ms of audio. Let R_{avg} be the average data transmission rate of the access point. Note that 802.11b allows for multiple data rates and most implementations support data rates of 11Mbps, 5.5Mbps, 2Mbps and 1Mbps. The actual data rate of a client (or an access point sending to this client) depends on the signal-strength of the client from the access point as well as other factors such as interference. The signal-strength in turn depends on the distance between the client and the access point. In general, depending on the special distribution of clients R_{avg} varies from 1 Mbps to 11 Mbps.

Table I shows a refined view of the overhead per packet in bytes or microseconds also taking the average data rate into account. Let T_P be the time taken to transmit the VoIP payload of P bytes. Further let $T_{overhead}$ be the average overhead per

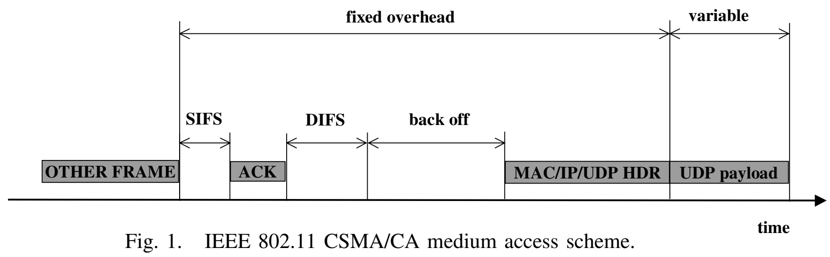

Fig. 1. IEEE 802.11 CSMA/CA medium access scheme.

Overhead	Bytes	Time (μs)
RTP Layer	12	$12 * 8/R_{avg}$
UDP Layer	8	$8 * 8/R_{avg}$
IP	20	$20 * 8/R_{avg}$
MAC	34	$34 * 8/R_{avg}$
Physical	24	192μs
SIFS		10μs
DIFS		50μs
ACK	14	$14 * 8/R_{avg}$

TABLE I

OVERHEAD PER SENT FRAME TAKING THE AVERAGE DATA RATE INTO ACCOUNT.

RTP packet. The channel throughput is simply given by

$$\frac{T_P}{T_P + T_{overhead}} \times R_{avg}.$$

While T_P is independent of the number of clients transmitting, $T_{overhead}$ varies with the number of clients. Specifically, $T_{overhead}$ can be divided into two components. First, the overhead incurred in transmitting the extra bytes of various networking layers and second, the overhead imposed by the Distributed Coordinated Function (DCF) of 802.11. Let T_{layers} and $T_{channel}$ denote these overheads respectively. Moreover, $T_{channel}$ comprises one SIFS interval, one DIFS interval, time to send an ACK and the average idle slots (T_{dcf}) per frame as seen on the channel. It is apparent that only T_{dcf} depends on the number of clients in the network. Moreover, as previous analysis studies have evaluated [1], [2], [7], [3], the dependence is non-linear. As the number of clients increases, the average idle slots per frame as seen on the channel decrease. For example, for a single client, assuming that the source always has a frame to send, there are 15.5 idle slots (i.e. 310μs) (Mean of uniform distribution $U[0, 31]$) per frame. If there are n clients, then the average number of idle slots are determined by two factors with opposing effects. First, after a successful frame transmission, the second frame is put on the channel by the client which calculated the minimum backoff among the n clients. However, it is also possible that more than one client calculated the same backoff resulting in a collision of the next frame or some subsequent frame, reducing the channel throughput. The collision probability increases with increase in n, counteracting the first effect. Further, upon collision, the contention window size is doubled, which leads to an increase in the average idle slots per frame. In effect, T_{dcf} is a concave function of n. If $n = 1$, however, there are no collisions and $T_{dcf} = 310\mu$s.

For VoIP connections, the access point receives frames from each of the n clients and sends VoIP packets to each of these clients. Assuming G711 a-Law codecs, the access point handles upstream traffic of $n \times 64$ kbps and sends $n \times 64$Kbps downstream traffic. The total channel throughput is therefore $n \times 128$Kbps. The maximum channel efficiency is given by

$$\frac{T_P}{T_P + T_{layers} + T_{SIFS+DIFS+ACK} + T_{dcf}}.$$

The maximum number of VoIP connections, n_{max}, all using G711 codecs is therefore given by

$$\left\lfloor \frac{T_P * R_{Avg}}{128000 \times (T_P + T_{layers} + T_{SIFS+DIFS+ACK} + T_{dcf})} \right\rfloor.$$

In our VoIP setup, each of the n clients has data to transmit very infrequently (every 10 ms for 80 byte payload), while the access point sends n times the data throughput downstream to the clients. Under this scenario, we argue that T_{dcf} can be approximated, even for large n, by the value of T_{dcf}, which results when an access point and a single other client always have a packet to transmit. In other words, from the perspective of the channel, the value of T_{dcf} is approximated by the value when there are exactly two active senders. The argument proceeds as follows: Assume that each of the n clients is using a 10 ms audio payload per packet. This implies an 80 byte RTP payload. The time to transmit this payload is 58 μs. In other words, with $R_{avg} = 11Mbps$, $T_P = 58$. The total overhead, not counting T_{dcf} is $T_{layers} + T_{SIFS+DIFS+ACK} = 500\mu$s. Assuming that the clients are not synchronized, the probability that a client, which is sending packets only every 10 ms, attempts to transmit a frame within 558μs of an ongoing transmission by another client is very low. In previous work on throughput analysis of 802.11 networks, T_{dcf} has been evaluated both via analytical approximations as well as via simulations. From these as well as from our own simulation, for the two sender scenario, the average back-off window size is determined to be 8.5 SLOTs and the collision probability is 0.03. Therefore, T_{dcf} for this case is given by $T_{dcf} = 8.5 \times 20 + T_W \times 0.03\ \mu s.$, where, $T_W = T_P + T_{layers} + T_{SIFS+DIFS+ACK}$.

A. Maximum number of VoIP connections and effect of codec selection

We use the above equations to compute the maximum number of VoIP connections a single 802.11b access point can support. These calculations are done for three standard codecs. ITU's G711 a-Law, G729 and G723.1. G711 allows 10 ms of audio data in 80 bytes of payload, without any compression. A single G711 VoIP stream therefore constitutes 64Kbps. G729 uses the same sampling rate of audio stream, but compresses the digitized data to pack 10 ms of audio in 10 bytes reducing the bandwidth of a single stream to 8Kbps. G723.1 further reduces the bandwidth requirement by compressing 30 ms of audio data into 24 bytes, which amounts to 6.4 Kbps per VoIP stream.

Table II tabulates the maximum number of VoIP connections for the three codecs. In these calculations, $R_{avg} = 11$ Mbps, which is the maximum possible transmission rate.

Audio (ms)	G711	G729	G723
10	6	7	
20	12	14	
30	17	21	21
40	21	28	
50	25	34	
60	28	41	42
70	31	47	
80	34	54	
90	36	60	61
100	39	66	

TABLE II

MAXIMUM NUMBER OF VOIP CONNECTIONS FOR DIFFERENT CODECS

The first column shows the audio payload sent in every RTP packet. For G.711 and G.729 codecs it is varied in increments of 10 milliseconds whereas for G.723 codec, it is varied in increments of 30 ms. The most important observation is the very low number of VoIP calls for lower payloads per RTP packet. For a payload of 10 ms audio, only 7 connections can be supported by a single access point. It is also worth noting that most commercial implementations of IP phones use a payload size of either 20 or 30 ms audio in each RTP packet. With G711 and 20 ms audio payload, the maximal number of VoIP connections is 12, and for 20 ms audio it is 17. In other words, the choice of payload size in IP phones results in very inefficient use of the available bandwidth in 802.11b. An obvious solution as shown by the numbers is to use larger payload per RTP packet. However, VoIP calls traverse both wireless and wired networks and the larger the payload, the worse are the delay, loss and jitter characteristics adversely affecting the VoIP call quality. Another fact worth observing is that for 20 and 30 ms payloads, choosing G729 or G723 codecs increases the maximum number of calls supported only by a small number. For instance, use of G729 over G711 with 20 ms payload only increases n_{max} by 2. This brings out a key limitation. While in wired networks (such as Ethernet), choice of codecs is highly effective in dealing with network load, in 802.11 networks, this is not the case. Comparing G729 and G723, even for higher payload sizes, n_{max} is almost equal. Therefore G729 should always be the preferred codec as it uses less compression.

As seen above, the total available payload bandwidth at small packet sizes is exhausted even for a small number VoIP of calls. An interesting observation, which is explored in detail in [6], is that the delay and jitter characteristics of the VoIP traffic are very much in the acceptable ranges even as this maximum number is crossed. In other words, the dominant factor that makes DCF unsuitable for VoIP is not that delay and jitter values are high, but that the effective available bandwidth is too low.

Another interesting fact to note is that for VoIP with small packet sizes, and because of the DCF behavior, delay and jitter values for these packets do not become unacceptable.

B. Effect of spatial distribution of clients in 802.11b

802.11b allows support for multiple data-rates and most commercial implementations support data transmission rates of 11Mbps, 5.5Mbps, 2Mbps and 1Mbps. The actual transmission rate of a client depends on the strength of the radio signal to/from the access point and other factors such as interference. For lower observed signal strengths, a client will transmit at a lower transmission rate to reduce the bit error rate. Since signal strength depends on the distance, R_{avg} depends on the spatial distribution of the clients. As the actual function is vendor dependent and R_{avg} is influenced by other factors such as interference, we study the effect on n_{max} for the whole range of R_{avg}.

Figure 2 shows n_{max} plotted against R_{avg}, where R_{avg} is varied between 1Mbps and 11Mbps for G711 and G729 codecs. The scale of the vertical axis in both is same for comparison purposes. The thick lines in each represent the typical payload implementation in commercial IP phones. The first obvious observation is the significant reduction in n_{max} for all payload sizes for lower values of R_{avg}. For instance, with a payload of 30 ms, n_{max} reduces from 17 at 11Mbps to only 4 at 1Mbps. This implies that the physical location of an access point in enterprises and public hot-spots is crucial, at least as far as supporting VoIP connections are concerned.

C. Maximum VoIP connections in 802.11a

The maximum data-rate in 802.11a is 54Mbps, approximately 5 times that of 802.11b. The value of the overhead T_{layers} of transmitting a total of 74 bytes for MAC, IP, UDP and RTP headers is therefore reduced proportionally. The physical layer overhead (PLCP header and preamble) is 24μs compared to 192μs for 802.11b. Further $SIFS = 16\mu$s and each $SLOT = 9\mu$s. T_{dcf} can be approximated by using exactly the same arguments as for the 802.11b case. In 802.11a, each client chooses a random backoff from a uniform distribution $U[0, 15]$ as opposed to the upper limit of 31 for 802.11b for successful transmissions. The average backoff when two 802.11 clients always contend for the medium is shown via simulations to be 41μs (4.5 SLOTs). The collision probability however is double that in 802.11b to 0.06. Therefore,

$$T_{dcf} = 4.5 \times 9 + T_W \times 0.06 \ \ \mu s.$$

Using the maximum value of $R_{avg} = 54$ Mbps, Table III tabulates the maximum number of VoIP calls that a single 802.11a access point can support.

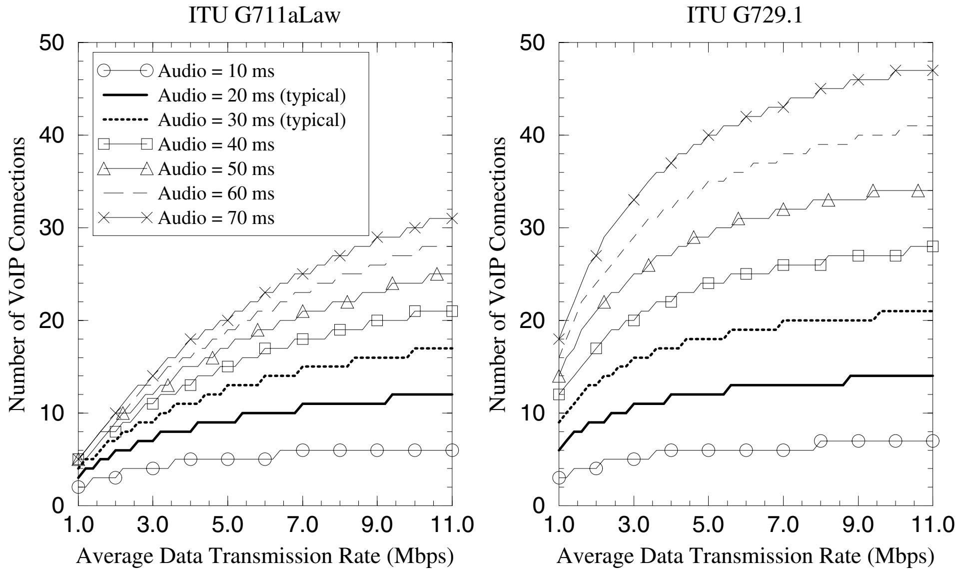

Fig. 2. Effect of spatial distribution/Load on the number of VoIP connections

Audio (ms)	G711	G729	G723
10	30	32	
20	56	64	
30	79	95	96
40	98	126	
50	116	156	
60	131	185	187
70	145	214	
80	158	243	
90	169	270	275
100	180	298	

TABLE III

MAXIMUM NUMBER OF VOIP CONNECTIONS FOR DIFFERENT CODECS WITH 802.11A (54 MBPS)

As is apparent by comparing these numbers from those of Figure II, n_{max} is approximately 5 times higher. The reason is simply the proportional increase in bandwidth and approximately the same channel efficiency.

IV. CONCLUSIONS

In this paper, we studied the behavior of VoIP over 802.11 networks, from the perspective of number of connections that a single base-station can support. An important conclusion is that given the choice of payload sizes in IP phones, 802.11b base-stations prove to be inadequate in handling large number of VoIP calls. In fact, the inherent channel inefficiency of 802.11b, at smaller frames sizes, limits the maximum number of VoIP calls to a very low number. The actual number of VoIP calls are further reduced by factors such as spatial distribution of the clients.

From the study, the use of larger payload per frame becomes apparent as the solution to increasing the number of VoIP calls. However, this needs to be balanced against the adverse effect of larger payloads on the multi-hop wired network in terms of delay, jitter and loss. This naturally lends to an optimization problem which is not addressed in this paper. Another conclusion drawn from the study was that the effect of codec selection does not help much in 802.11 networks, which is in contrast with wired Ethernet.

REFERENCES

[1] G. Anastasi and L. Lenzini. QoS provided by the IEEE 802.11 wireless LAN to advanced data applications: a simulation study, In *Wireless Networks*, Vol. 6, pp 99–108, J.C. Baltzer AG, Science Publishers, 2000.

[2] G. Bianchi. Performance Evaluation of the IEEE 802.11 Distributed Coordination Function. In *IEEE Journal on Selected Areas in Communication*, Vol. 18, No. 3, March 2000, pp. 535–547.

[3] H. S. Chayya and S. Gupta. Performance modeling of asynchronous data transfer methods of IEEE 802.11 MAC protocol. In *Wireless Networks* Vol 3, 1997, pp. 217-234.

[4] b. P. Crow, I. Widjaja, J. G. Kim and P. Sakai, "Investigation of the IEEE 802.11 medium access control (MAC) sublayer functions," In *Proceedings of INFOCOM'97*, volume 34, pages 126-133, April 1997.

[5] P. Fergusan and G. Huston *Quality of Service*, Wiley Computer Publishing, John Wiley and Sons, ISBN 0-471-24358-2, 1998.

[6] S. Garg, M. Kappes, "An experimental study of throughput for UDP and VoIP traffic in IEEE 802.11b Networks", *Submitted for publication*

[7] A. Heindl and R. German. Performance modeling of IEEE 802.11 wireless LANs with stochastic Petri nets. *Performance Evaluation*, 44 (2001), 139-164.

[8] J. Schiller. *Mobile Communications*, Addison Wesley, ISBN 0-201-39836-2, 2000.

[9] J. L. Sobrinho and A.S. Krishnakumar "Real-Time traffic over the IEEE 802.11 MAC protocol", In *Bell Labs Technical Journal*, Autumn 1996.

[10] M. Veeraraghavan, N. Cocker and T. Moors "Support of voice services in IEEE 802.11 wireless LANs", In *Proceedings of INFOCOM'01*, 2001.

[11] J.-Y. Yeh and C. Chen "Support of multimedia services with the IEEE 802.11 MAC protocol." In *Proceedings of ICC'02*, New York, April 2002.

[12] ISO/IEC and IEEE Draft International Standards, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," ISO/IEC 8802-11, IEEE P801.11/D10, January 1999.

[13] http://grouper.ieee.org/groups/802/11/