Voice over an OFDM based Wireless LAN

Barry M. G. Cheetham , Abdulmola I. Eshhubi
Department of Computer Science
University of Manchester
Manchester M13 9PL
UK

Email: barry@ man.ac.uk

Abstract

This paper addresses the problem of efficiently integrating speech communication devices, such as cordless telephones, with an "orthogonal frequency division multiplexing" (OFDM) based wireless LAN. It is argued that traditional "voice over IP" (VoIP) approaches are inefficient for wireless LANs in terms of system overheads, and that more recent proposals, such as the "5-UP" protocol, offer greater promise with the co-existence of OFDM, for high speech packetised data, and simpler (e.g. single carrier) modulation schemes for lower data-rate devices with real time quality of service requirements. A new way of supporting packetised voice is suggested here which is compatible with the IEEE 802.11a and 802.11g standards in "DCF" mode and the "EDCF" features being proposed for 802.11e. The intention is to avoid the "5-up" MAC approach which requires the use of the "PCF" beacon to time-share the different protocols and to preserve full OFDM functionality for high-speed data. The new approach will operate alongside 802.11a or 802.11g systems, being compatible with both the physical and MAC layers. At the physical layer, the packetised speech communication will be based on single carrier (SC) modulation. The speech devices must be compatible with the CSMA collision avoidance protocol, to be capable of detecting whether other OFDM or SC devices are active, but need not be capable of demodulating or transmitting OFDM. For speech communication the new system's efficiency will be much higher than VOIP because the speech channels will be frequency division multiplexed. The power efficiency of the speech units can also be much higher than if they used full OFDM. Some of the advantages and problems remaining with this approach will be discussed.

1. Introduction

Wireless networks are an alternative to high cost wired data communication systems in new broadband applications and can also provide conventional telephony and computer networking in environments, and in parts of the world, where a suitable wired infrastructure does not exist. Traditional broadcasting and the first generation of mobile telephones use analogue modulation techniques with frequency division multiple access (FDMA). Later generations of mobile telephony networks are based on digital modulation techniques using combinations of FDMA and either 'time division multiple access' (TDMA) or 'code division multiple access' (CDMA). In recent years, a more sophisticated modulation technique known as orthogonal frequency division multiplexing (OFDM) has received much attention [1]. This technique is considered suitable for achieving high bit-rates in wireless environments subject to multipath propagation and consequent channel fading. OFDM is already being used for broadcasting (DAB and DVB-T), "ADSL" modems and wireless LANs but not for widely for voice telephony. This work will focus on the use of OFDM with wireless LANs, which is supported currently by two standards, IEEE 802.11a and its European equivalent

HIPERLAN2. A recently approved standard, IEEE 802.11g, incorporates an OFDM based scheme designed to operate alongside IEEE802.11b in the 2.4 GHz band. It is claimed that with OFDM, IEEE 802.11g will enable data transmission speeds of up to 54 Mbps with backwards compatibility to 802.11b infrastructures [2]. There is great interest in integrating wireless LANs and wireless telephone networks so that communication takes place over a single allocated frequency band. This paper addresses the problem of efficiently integrating speech communication devices, such as cordless telephones, with an OFDM based wireless LAN.

2.The "DECT" Standard for Cordless Telephones

This standard for "digital European cordless telephones" operating in the 1880 to 1900 MHz band was approved by ETSI in July 1992. Its main purpose is to provide mobility within private branch (PABX) telephone exchanges [3] allowing communication between a base-station and up to about 120 cordless phones at ranges of up to a few hundred metres. DECT has ten carriers each being used for 24 TDMA speech channels and thus accommodating twelve two-way conversations. Each speech channel segments its data in 10 ms frames and is allocated 480 bits per frame. With 100 frames per second for each of the 24 frames, a total bit-rate of 1152kb/s must be conveyed by each of the ten carriers. With binary (Gaussian MSK) signaling and a bandwidth efficiency of 0.75 bits/Hz, a bandwidth of 1.728 Hz is sufficient for each of the ten carriers. DECT is strictly configured according to the OSI standard with "control" and "user" planes above the physical and MAC layers described above. The speech is encoded by 32kb/s ADPCM (G721) which therefore requires 320 bits per 10ms frame. The remaining 140 bits in each 480 bit TDMA slot are used for a 32 bit preamble, 40 guard bits, and a 64-bit "A-field" comprising an 8-bit "header", 40 "data" bits and 16 "CRC" bits.

3.Orthogonal Frequency Division Multiplexing

OFDM is a multiple carrier scheme where the data being transmitted is carried on a large number of independent sinusoidal waveforms whose frequencies span a bandwidth much larger than the "coherence bandwidth" of the radio channel. This means that multipath propagation causes 'frequency selective fading' i.e. reductions in signal strength which vary significantly with frequency across the band. With single carrier digital modulation schemes, as used for second generation mobile telephony for example, some form of equalisation is required to combat the inter-symbol interference (ISI) that would result from frequency selective fading. Such equalisation would generally use a time-domain adaptive filter, which introduces considerable complexity. The multicarrier nature of OFDM combats the effect of frequency selective fading by frequency-domain processing which is much more flexible and efficient. A key feature of OFDM is the transmission of blocks of data separated by "guard-bands" constructed in a particular way. The blocks of data result from the application of the inverse 'fast Fourier transform' (FFT) to achieve the modulated signal. The guard-bands are "cyclic prefixes" obtained from the cyclic nature of the FFT. The cyclic prefixes have several crucial advantages such as the ability to achieve synchronisation, avoid ISI from one block to the next, and, most importantly of all, to allow the frequency-domain processing needed to compensate for frequency selective fading to be carried out in the FFT frequency-domain where cyclic convolution rather than linear convolution describes the effect of the selective fading on the channel.

4.IEEE 802.11 Standards

The IEEE 802.11a standard [7] for wireless network communication at data rates up to 54 Mb/s, in 20 MHz frequency bands around 5 GHz, is based on OFDM as described above. Each 20 MHz band is spanned by 64 separate frequencies, at 312.5 kHz separation, of which 48 carry data by PSK or QAM modulation. IEEE 802.11b employs a single carrier modulation technique known as complimentary code keying (CCK) for the preamble/header and the payload for each packet. IEEE 802.11g uses the same CCK scheme for preamble /header and is able to use OFDM at a higher bit-rate for the payload. One of the features of the IEEE 802.11 range of standards is the ability to work in "distributed control function (DCF)" mode without being under the control of a central controller. This is possible with the use of a “CSMA/CA” collision avoidance protocol similar to that used by wired Ethernet [4]. A RTS/CTS (request to send / clear to send) signalling mechanism between devices is included in the IEEE802.11 standards to combat the “hidden node” problem [5] where two devices attempting to communicate with a given node are out of range with each other. An alternative “point control function (PCF)” mode is also available where a base station centrally controls all activity within its sphere of influence. IEEE 802.11 standards are likely to become predominant for wireless LAN operations.

A new version, IEEE802.11e [11] is still being defined with the aim of enhancing the medium access mechanism to provide a means of giving priority to some types of frames, such as frames carrying speech or multimedia transmissions, which may be more time-critical than others. Although the details have not been finalized, it is likely that a new access mechanism based on a “hybrid coordination function” (HCF) and an “enhanced DCF” (EDCF) will be introduced based upon an idea by D.J.Cheng and R.S. Chang [13] and similar to a mechanism used in Hiperlan. The enhanced DCF mode differentiates between different priorities by varying the size of the "minimum contention window" giving lower minimum delays and therefore faster medium access to some devices. Also, the concept of inter-frame spacing (IFS) , with DIFS, SIFS and PIFS as conventionally used in contention mode IEEE802.11, is extended to include an "arbitration inter-frame spacing" (AIFS) which is variable and introduces differential delays for different devices wishing to access the medium. Again devices with lower delays have priority. "Packet bursting" is also being proposed to allow devices to send more than one frame having gained access to the medium.

Due to demands for the different networks to converge under one type of network, there is great interest in integrating wireless LANs and mobile telephone networks so that communication take place over a single allocated frequency band in the 2.4 GHz or the 5 GHz bands. WLANs and speech terminals both operating within the same band should be able to communicate and the maintenance of a single radio network rather than two separate ones would be more economic, for example with respect to the cost to a commercial company of access points and external connections.

5.Voice over IP (VoIP) and VoWLAN

Integration between speech terminals and computer networks can be achieved without affecting the physical layer using voice over IP (VoIP). Commercial systems already exists using this approach. Voice over IP (VoIP) is a term used for the packetization of continuous real time speech and the transmission of speech by the mechanisms (IP) devised for data over computer networks and/or the Internet. The normal overheads of IP packets, as encapsulated in IEEE802.11 packets, can be reduced for the confines of a wireless LAN thus economising on the required bit-rate. However, the use of VoIP techniques adapted to wireless LANs, VoWLAN say, is innefficient and, according to certain manufacturers [6], limits the number of simultaneous real time speech channels to about eight with IEEE 802.11b. VoWLAN can not guarantee a specified quality-of-service (QoS) because of the connection-less nature of most computer networks and the inability of such networks to guarantee error-free delivery of packetized data within a specified time frame. Therefore quality control mechanisms such as RTP/RTCP [4] must be introduced along with lost or delayed packet concealment techniques. A further problem with VoWLAN is that the IEEE 802.11 preamble/header and packetization overheads can make the WLAN very inefficient for the relatively short but regular packets needed for speech communications.

To illustrate this point for IEEE 802.11b, if we have 40 ms speech packets, each encapsulating four 480-bit (10ms) DECT frames, assuming 200 pre-amble/header bits at 1Mb/s and 1920 payload bits at 11 Mb/s, we require 200+175 = 375 μs per packet. With 25 frames per second and two-way transmission per user, this requires 375 x 25 x 2 μs ≈ 19 ms per second, i.e. 1.9 % of total network capacity, for each user. This neglects the MAC overheads and any encapsulated IP headers but probably overestimates the DECT control information needed in the payload. With ten voice users, the cost would be 10 x 1.9 % i.e. 19 % of network capacity which is considerable. This loading could reduce to about 1.2 % of network capacity per user with IEEE802.11g and 54Mb/s OFDM payload, and could be made still less with IEEE80211a because of the more efficient pre-amble/header transmission. However, with VoIP, each speech terminal would require a full IEEE 802.11 capability which for IEEE 802.11a or IEEE 802.11g would need a complex and power consuming chip-set incorporating an FFT. The linear nature of OFDM though suitable for portable computers with large batteries is not so suitable for mobile phones because of the power inefficiency of the linear amplifiers that are required for OFDM transmission. OFDM based mobile or cordless phones would be inefficient in terms of power consumption. A solution to these problems is given as an even newer proposal (not yet adopted as a standard) known as the 5 GHz Unified Protocol ("5-UP") [8].

6. "5-Up"

This proposal by Atheros Communications Inc. [8] is an extension to IEEE 802.11a as a 5GHz WLAN standard to support data transfer rates to over 54 Mbps and also a wide variety of lower-power lower-speed devices carrying diverse traffic types to coexist and interoperate within the same unified wireless network. The principle of '5-UP' is to allocate some of the OFDM carriers on an individual basis to other lower speed devices, which could be, for example, cordless phones with single carrier modulation (SC) schemes. Data devices must avoid using the "SC" carriers when they are being used for "5-up", and all SC communication takes place between a central access point (AP) and an SC terminal. The frequencies used by the SC terminals must be locked to the transmissions from the AP to an accuracy of about 1-2 parts per million (≈ ±10 kHz) to avoid co-channel interference. Frequency correction may be applied to all carriers as received at the AP. To allow the geographically distributed SC transmissions to be demodulated by a single FFT process per quasi-OFDM symbol all concurrent SC symbols must arrive within the appropriate 4 μs symbol slot. Misalignment in the time of arrival of each SC symbol will be less than about 100 ns for a typical WLAN environment and therefore the "cyclic extension" guard time of 800ns is more than sufficient to allow for this as well as multipath reflections. Power control is needed to avoid SC transmissions saturating the receiver and deviations from precise orthogonality causing excessive interference between OFDM carriers. Perhaps surprisingly, this can be achieved without unreasonable complexity [8]. Frequency hopping and antenna diversity may be used to reduce the effect of multi-path fading on the SC channels. Although 5-UP has interesting features, it has the disadvantage that restrictions are required to the operations of the basic 802.11a and 802.11g standards to allow 5-UP operation to co-exist. The solution proposed for the interoperability of 802.11a or 802.11g with 5-UP requires a degree of central control, using the "point coordination function" (PCF), where the available time is divided between standard 802.11a or 802.11g and 5-UP operation. The “5-up” MAC uses PCF, transmits “5-up” beacons on each carrier and sends information about frequency hopping etc. The use of the PCF may be considered an undesirable feature of “5-UP” as many studies [e.g. 10] report the PCF to be an inefficient way of integrating data and interactive voice communication within wireless LANs.

7.Other Approaches

Many voice over wireless LAN techniques use the point co-ordination function (PCF) [9]. Liu and Wu [9] propose a novel distributed co-ordination function (DCF) mode approach which uses the IEEE "beacons" and power saving facility (MAC approach) to achieve a "pseudo-TDM" technique for each speech channel. BP Crow & al. [10] and many other researchers conclude that PCF poorly supports voice. Many researchers [e.g.11-13] are now looking at the problem of integrating voice with contention mode wireless networks and exploring the new MAC sub-layer of the proposed IEEE802.11e standard. Popular MAC sub-layer approaches, described in [11], include “distributed fair scheduling” (DFS) and “blackburst”. There appears to have been little follow-up from the “5-UP” physical layer proposal [8] so far.

8.Our Proposal

We now propose a new way of efficiently integrating a cordless phone system with an OFDM based wireless LAN which is compatible with the standards 802.11a and 802.11g without the need for central control. This proposal will operate alongside 802.11a or 802.11g systems, being compatible with both the physical (PHY) layer and multiple access control (MAC) sub-layer. The speech communication will be packetised and, at the physical layer, based on single carrier modulation within the 5 GHz or 2.4 GHz bands. The speech units will be capable of detecting whether other high data rate OFDM or low data-rate devices are accessing the radio medium and of sending/receiving RTS/CTS signals according to the relevant standard. However, they need not be capable of transmitting or receiving high speed OFDM data packets. The MAC sub-layer, which manages the access to the medium, employs "carrier sense multiple access with collision avoidance" (CSMA-CA or DCF), i.e. the same mechanism as 802.11a and 802.11g.

For speech communications, the new system's efficiency will be higher than VoIP based systems because the speech channels will be multiplexed; i.e. when one speech unit is transmitting its payload, other speech units will be allowed to transmit simultaneously at 250 kbaud using different carrier frequencies selected from the set of frequencies used by OFDM. Each speech unit employs single carrier (SC) modulation, with 4μs symbol timing which includes the 800ns cyclic prefix (C.P) to be compatible with IEEE 802.11 OFDM. The advantages of having such integration are that the new communication system will not interfere with existing IEEE 802.11a, b or g based WLAN devices and the number of simultaneous calls will increase over what is possible with VoIP. Also because of the single carrier modulation approach, the speech units will be efficient in terms of complexity and power consumption. Frequency-hopping can occur across all 48 data carriers now and should therefore be simpler and more effective than with "5-up". When less than 48 speech channels are needed, frequency usage can be made even more efficient in many obvious ways.

It is interesting to consider the differences between a normal OFDM system and our "distributed ODFM" transmissions. With normal OFDM, the modulation is achieved numerically by the FFT, and although an array of carriers is effectively modulated individually by QPSK or QAM, these modulated carriers are never seen separately. The required amplitudes and phases for each 4us symbol are produced as complex numbers which are then applied to the inverse FFT algorithm. This produces 64 time-domain samples which are cyclically extended to 80 samples. This may be thought of as simulating an array of SC transmitters whose carrier frequencies are exactly synchronised in frequency and phase and where the modulation at 250kbaud is by symbols that are rectangularly windowed in time, remaining constant for the whole 4us symbol time. The symbols are therefore not band-limited as they would normally be in a SC transmitter. There is (almost) no pulse-shaping and the transmission from each simulated SC modulated carrier in principle occupies the whole 20 MHz band, though with zeroes in the spectrum coinciding with the frequencies of other carriers.

With a distributed OFDM system with all SC terminals transmitting to the AP having synchronised their carrier frequencies to an AP transmission, we get the transmissions (normalised in maximum power to within ±3 dB of some maximum (as in "5-up")) without phase synchronisation. Some of the transmissions will suffer fading and be received at low power levels. Unlike normal OFDM, the fading will be different for each carrier because the transmitters are in different places geographically. So we cannot assume a delay spread which allows the necessary equalisation to be calculated say for every other carrier, or one carrier frequency in five (using interpolation for the ones between). All carriers can suffer quite different degrees of fading compared with that suffered by neighbouring carriers as the transmitters are not at the same point in the room.. However this presents an opportunity as we could select the frequency for each SC device intelligently such that it does not incur fading, where a different frequency at that location would suffer fading. This may be a viable alternative to frequency-hopping making better use of all frequencies.

Although the use of PCF mode is not required, the role of a central access point (AP) is to provide access to outside PSTN and network links and the timing synchronisation needed for compatibility and intercommunication between the new SC terminals and standard IEEE 802.11 devices. The AP will transmit to all SC devices simultaneously as if they were a single standard OFDM device. If precise enough carrier and frame synchronisation is achieved, the transmissions received from all SC devices may be demodulated by a single FFT as if they came from a single OFDM device. The AP can retransmit the speech it receives across the WAN to another user or onto an outside link. Clearly the AP or any other computer with normal IEEE802.11 access to the wireless LAN can act as a cordless phone. SC devices can send data to the access point only and cannot receive normal OFDM transmissions. SC devices must respect all IEEE 802.11 conventions even those it does not make use of (e.g. the PCF).

Estimation of Parameters

Investigations are being carried out to evaluate the potential of this new approach and to determine the best way to overcome the problems which will inevitably degrade its performance, such as channel error, packet loss and packet delay. Since speech must be packetized with the new system, investigations must also be carried out to determine the best way of doing this. With IEEE802.11a and g, OFDM symbols are transmitted in 4 μs giving a maximum throughput of 250 k symbols/second. Each symbol can carry up to 6 bits per carrier (using 64-QAM) this rate being normally reduced to 4.5 bits per carrier by a ¾ rate convolutional coding scheme. As there are 48 carrier capable of carrying data, this gives a maximum bit-rate of 48 x 4.5 x 250 kb/s = 54 Mb/s. The distances over which this bit-rate is achievable in practice will be restricted by transmission loss and interference, and lower bit-rates (48, 36, 24, 18, 12, 9 and 6Mb/s) are available. The two lowest bit-rates, i.e. 9 Mb/s and 6Mb/s, use binary phase shift keying (BPSK) and 3/4 or 1/2 rate convolutional coding to achieve 48 x (3/4) x 250kb/s = 9 kb/s or 48 x (1/2) x 250 kb/s = 6 Mb/s. To achieve 18 Mb/s and 12 Mb/s quadrature phase-shift keying (QPSK) is used on each of the 48 data carriers. If we wish to allocate part of the WLAN network capacity to up to 48 speech channels each derived from DECT devices with modified radio interfaces, we may, in the first instance, choose to simply concatenate four consecutive 10-ms DECT frames, with all their preambles and headers etc. to form a 40 ms frame requiring 4 x 480 bits = 1920 bits. At 25 frames/sec this would requite a bit-rate of almost 50 Mb/s. If this were achieved for each speech channel by applying QPSK modulation to one of the 48 OFDM carriers, the cost to the WLAN would be 10ms out of every second, which is 10% of the total network capacity, plus the additional cost of IEEE802.11a/g control information per packet (preamble/header etc.). We must also estimate the cost of the RTS/CLS protocol, if this is used. With 802.11g the preamble/header overhead is about 200 bits at 1 MHz and if there are 25 frames per second as proposed, the total additional cost is only 200 x 25 x 10-6 seconds per second = 5 ms per second which is 0.5 % of network capacity. If we add to this an 1000 x 25 x 10-6 seconds per second overhead for RTS/CTS (including associated MAC delays) we conclude that total of 13 % loss of capacity for data is incurred by this simple voice over WLAN scheme. Many economies are possible, and some of these are probably needed. For example we could reduce the speech bit-rate by compression (to say 8 kb/s) , increase the packet size from 40 ms to 100 ms, and use a higher level modulation scheme, say 64-QAM instead of QPSK. A better way of defining the 10ms speech + DECT control packets (better than simply concatenating DECT frames) would also be beneficial. This would reduce the overhead on the network to less than 1%, ignoring the normal costs of MAC protocols, though there are some disadvantages (e.g. delay) and increases in complexity with these modifications. A balance between the two extremes mentioned here is likely to be chosen, and an adaptive scheme would probably be even better.

Quality assurance

Absolute QoS guarantees are not possible, but there are many reasons to believe that this system can be more efficient than those which try to offer such guarantees using the PCF facility. The use of the HCF and EDCF facility as being proposed for IEEE802.11e can improve the QoS. Nevertheless, delayed/lost packets concealment techniques will be needed. In contrast with wired LANs, damaged or lost packets will be more of a problem than delayed packets except when the delays are due to VoIP over the Internet.

9.Conclusions

We propose a method for integrating wireless telephony with OFDM based wireless LANs conforming to the IEEE 802.11a and 802.11g standards. Unlike "5-up", this method does not rely of the availability of a point co-ordination function (PCF) to allocate separate time slots to non-OFDM devices. It is compatible with the contention modes of IEEE 802.11a, g and can benefit from the HCF and EDCF functions being proposed for IEEE 802.11e. The need to reserve some OFDM frequencies for single carrier devices, thus restricting normal OFDM operation between standard IEEE 802.11a/g data terminals is also eliminated. This introduces the need for efficient packetised speech transmission and provision for minimising the effect of corrupted, delayed or lost packets. Although there is much detail to work out, this appears to be a promising approach.

10.Reference

[1]J. H. Stott, "Explaining some of the magic of COFDM", Proceeding of 20th International Television Symposium 1997-Paper 28/08/98 http://www.bbc.co.uk/rd/pubs/paper_15.htm I.E. Sutherland, "Micropipelines",*Communications of the ACM*, Vol.32, No.6, pp.720-738, June,1980.

[2]J. Zyren, "IEEE 802.11g Explained", Wireless Networking, December 2001.

[3]T.S. Rappaport, "Wireless Communications", Prentice-Hall, 2002

[4]A.S. Tanenbaum, "Computer Networks" 4th ed, Prentice Hall, 2003

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

[6]Spectralink VoIP product: http://www.spectraLink.com /products/pdfs/NetLink_FAQ.pdf

[7]M.S. Gast "802.11 Wireless Networks", O'Reilly &Ass. Inc., 2002, ISBN 0-596-00183-5

[8]B. McFarland, G. Chesson, C. Temme & T. Meng, "The 5-UP Unified Protocol for Multiservice Wireless Networks", IEEE Communications Magazine, November 2001, vol.39 no.11, pp. 74 - 80. (Atheros Comms Inc.)

[9]HH Liu & JCL Wu "A scheme for supporting voice over IEEE802.11 WLAN", Proc. Natl Asci Council, ROC(A), vol 25, no. 4, 2001, pp 259-268

[10]BP Crow & al. "802.11 Wireless Lans", IEEE Comms Mag, Sept '97, pp.116-126

[11]A. Lindgreen, A. Almquist & O. Schelen, "Evaluation of Quality of Service Schemes for IEEE 802.11 wireless LANs", Proc 26th Annual IEEE Conf. on Local Computer Networks (LCN 2001), Nov 2001

[12]A. Kopsel & A. Wolisz, "Voice transmission in an IEEE 802.11 WLAN based access network", Proc Int. Workshop on Wireless Mobile Multimedia, Rome, Italy, 2001

[13]D.J. Cheng & R.S. Chang "A priority scheme for IEEE802.11 DCF access method". IEICE Transactions on Communications, E82-B(1), January 1999