Professional Quality Voice over WLAN

Marc Kuhn[1], Azadeh Ettefagh[1], Michael Kuhn, Andrew Lunn[2], Barry Cheetham[3], Marjan Spegel[4]
[1]Swiss Federal Institute of Technology (ETH) Zurich, Communication Technology Laboratory, CH-8092 Zurich, Switzerland; email: kuhn@nari.ee.ethz.ch
[2]Ascom Systec AG, Gewerbepark, Mägenwil, CH-5506 , Switzerland, andrew.lunn@ascom.ch
[3]University of Manchester, Dept. of Computer Science, Oxford Rd, Manchester, UK, barry@man.ac.uk
[4]Jozef Stefan Institute, Dept of Intelligent Systems, Jamova 39, 1001 Ljubljana, Slovenia, marjan.spegel@ijs.si

Abstract - **A key element of generalized network access is the convergence of different services within one system. Future devices and applications will integrate professional quality real time voice and video traffic with packetized data. This will replace the current solution, where in-house wireless networks have two parallel infrastructures: the telephony infrastructure providing low-bandwidth but high quality of service, and the computer network for high data-rate but best-effort bursty traffic.**

In this paper we present a novel approach to merge both wireless networks by integrating professional quality telephony into WLANs to get high quality "voice over WLAN". This will lead to an innovative application in the field of wireless and mobile communications. The approach is to merge the upper layers of the DECT protocol stack with a current WLAN physical layer and MAC layer technology, using a protocol adaptation layer (PAL).

I. INTRODUCTION

Voice over IP (VoIP) is becoming a very important technology for many applications including low quality long distance Internet telephony, small office telephony systems and niche applications where data networks are dominant. Over the Internet and busy links shared with data, VoIP is likely to have unacceptably large delay and/or unreliable packet delivery within the required time period. Acceptable performance is usually obtained in dedicated packet switched networks offering large and reliably available bandwidth (e.g. [9]). Therefore VoIP use is limited in many new access networks - especially wireless ones - for reasons related to resource management, resource cost and QoS control. Currently available products for "voice over WLAN" are based on VoIP and have additional drawbacks, e.g. high power consumption, small coverage, poor (or even no) handover provisions between base stations etc.

In this paper we present a new approach to "voice over WLAN". It is based on a combination of a digital cordless telephony standard and current wireless LAN (WLAN) technologies, and does not use VoIP technology. This approach has been developed within the WINDECT (wireless local area network with integration of professional-quality DECT telephony) project [5]. The aim of the WINDECT project is to enable high quality of service for voice over WLAN by integrating professional-

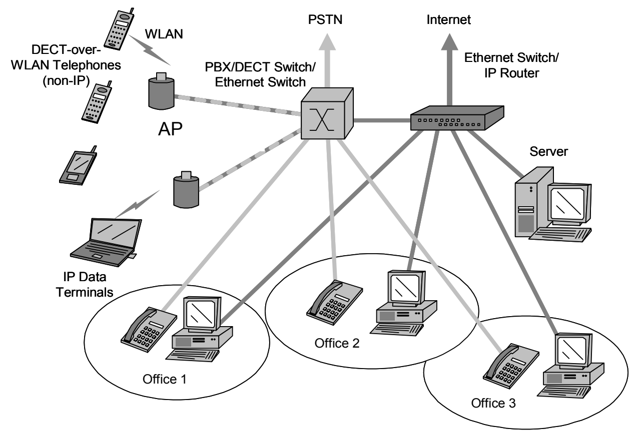

Fig.1. Converged in-house network with DECT over WLAN

quality DECT (digital enhanced cordless telecommunications) telephony [4] into WLAN technology, whereby crucially the quality of service in the telephony part is not impaired; in addition the compatibility to existing WLAN data modems is guaranteed. In the WINDECT approach the lower layers (PHY, MAC) conform to current WLAN standards, while the upper layers are defined according to DECT. In between a "Protocol Adaptation Layer" (PAL) is used as an interface. This approach promises many advantages over VoIP because the upper layers of DECT are optimized for the voice application.

A. Advantages of the WINDECT approach

VoIP introduces much more overhead than a solution based on DECT. VoIP headers allow that the voice packets can be routed through an IP network. DECT, on the other hand, simply passes the packet to the PBX (Private Branch eXchange). The PBX then decides what to do with the packet. In the scenario we are considering in this paper, the APs are directly attached to the PBX (see Fig. 1); therefore IP/UDP/RTP headers are mostly redundant. Furthermore VoIP is aimed at more heavyweight implementations; due to the large overhead a powerful embedded system is required. DECT is a much simpler standard in its GAP [4] form. This allows it to be implemented on lightweight devices, with minimal processor, RAM and FLASH requirements. This is then reflected in the battery requirements, standby and call times and finally leads to lower costs.

II. DECT: DLC AND NWK LAYER

DECT [4] is a low-power two-way digital wireless communication system. It is the standard for digital cordless telephony, designed for wireless use and fully integrated with the PSTN (Public Switched Telephone Network). It is a well-established technology for professional-quality telephony, and applied worldwide (the North American Personal Wireless Telecommunications standards PWT and PWT/E (TIA) are based on DECT). DECT uses 10 ms ADPCM (adaptive differential pulse code modulation) frames (32 kb/s). ADPCM is a ‘wave-form’ coding technique of moderate complexity which incurs virtually no delay additional to any framing delay. Standard 64 kb/s digitized speech (A-law PCM) is compressed into 32 kb/s ADPCM form with little or no discernable loss of quality by exploiting the correlation between successive speech samples.

The data link control (DLC) layer provides reliable data links to the network (NWK) layer. In the DECT standard the DLC layer is divided into the C-plane and the U-plane, which deal with control signalling and user data transport, respectively (a detailed definition of the DECT DLC layer can be found in [4]).

The network (NWK) layer of DECT organizes the information exchange: it uses the radio link on behalf of an application. The basic set of messages exchanged between NWK layer peer entities supports the establishment, maintenance and release of calls; additional messages support a range of extended capabilities (for details to the DECT NWK layer see [4]).

Although providing high quality voice, DECT is highly integrated, i.e. component costs are very low.

III. WLAN: PHY AND MAC LAYER

The WINDECT approach for physical (PHY) layer and medium access control (MAC) layer is based on the WLAN standards IEEE 802.11a [6], IEEE 802.11 [1] with extensions 802.11e (QoS) [2] and 802.11h (Spectrum and Power Management) [7]. An access point (AP) acts as a DECT base station (actually as the PHY and MAC layers of a base station) as shown in Fig. 1 and supports voice channels and additional data channels.

The physical layer IEEE 802.11a uses OFDM with up to 54 Mb/s in the 5 GHz range.

The MAC layer conforms to IEEE 802.11 with extension 802.11e. The legacy IEEE 802.11 standard [1] contains two access schemes, the Distributed Coordination Function (DCF) and the Point Coordination Function (PCF). These schemes control how terminals gain access to the available bandwidth. So far only DCF has been widely available, but PCF is intended to provide better support for real-time services. The “e” group [2] will add a third and fourth scheme, the HCCA (HCF Controlled Channel Access) and EDCA (Enhanced Distributed Channel Access).

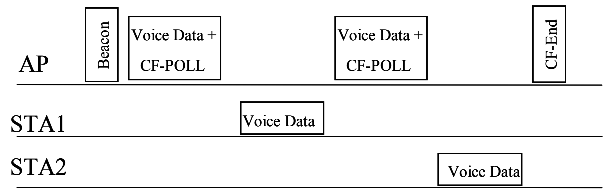

Fig. 2: HCCA based audio data transfer

The HCF (Hybrid Coordination Function) provides the required infrastructure to allow call admission control and bandwidth reservation. It adds the concept of a transport stream, with QoS parameters such as required bandwidth, delay and jitter. The AP can decide wheter to accept new streams given the current resource situation. Once a stream has been accepted it can then ensure that the requested bandwidth is available and that the delay and jitter constraints are fulfilled [3].

A. DECT TDM to WLAN HCF: Basic MAC scheme

DECT provides a high QoS by making use of a strict TDM (time division multiplex) structure [4]. This ensures that each station (STA) has a guaranteed time when it can transfer its voice packet from/to the RFP (Radio Fixed Part, in WLAN terminology: AP). To support something similar in an IEEE 802.11 WLAN, HCCA is used in the WINDECT approach for the voice traffic, EDCA for the data traffic. Each DECT AP will send out beacons at regular intervals. The beacons indicate the start of the Contention Free Period (CFP). For each STA that has an ongoing connection, the AP will transfer one voice data packet with a piggybacked CF-POLL message. The STA replies with its voice data packet. The AP then exchanges data with other STAs with data connections, as shown in Fig. 2. Once all STAs with active connections have been polled, the AP finishes the CF period by broadcasting a ‘CF-End’ message. In the following Contention Period (CP), data (regular data in contrast to voice data) traffic is transmitted.

By using HCCA in this way it can be arranged that each STA is able to transfer its voice data packet with a pre-defined QoS. The jitter is low since the beacons are sent at regular intervals. If the AP can be made to poll the STA in the same order every time, the jitter can be reduced further. Note that the WINDECT approach does not use ACKs. This means corrupt or lost packets will not be retransmitted. This is in keeping with DECT that also does not perform retries. It also helps to keep jitter to a minimum though packet loss concealment (PLC) must be implemented to reduce the perceived effect of packet damage or loss.

On average, for a duplex channel, at least one DECT packet containing 84 bytes payload (80 bytes of PCM audio data plus 4-bytes header) must be transferred every 10 ms in each direction (note that for WINDECT, PCM as well as ADPCM can be used). However, the WLAN packet does not have to correspond exactly to the standard DECT packet, as long as the average data rate and the jitter/delay requirements are met. If, e.g., the WLAN PHY works with 6 Mbit/s data rate, the transfer of one DECT duplex channel takes 348 µs incl. SIFS (Short Inter Frame Space), at a data rate of 54 Mbit/s it takes 136 µs.

The beacon intervals of the considered WLAN MAC are multiples of 1024 µs which eliminates the possibility of mapping the 10 ms DECT timing directly into the WLAN world. This means that a period of 10.24 ms or 20.48 ms must be used. When considering the power efficiency and handover issues to be discussed later, a period of 20.48 ms between voice data frames is used, whereas the beacon period is 10.24 ms. I.e., voice data of an active voice channel (containing about 20 ms of speech) is transmitted every other beacon. Half of the length of the beacon period, 5.12 ms, is the maximum length of the CFP, the remaining (≥ 5.12 ms) is used as CP. Data frames are used as variable-length DECT payloads to adapt between the 10 ms DECT and the 20.48 ms WLAN period. The audio data is considered as a stream from which as much data as available is transmitted in each period. DECT payloads will still have closely similar sizes, but they will contain a few PCM samples more or less as needed to reach the exact effective data rate used by the DECT protocol layers. In practice, this means that 99% of the packets transferred will contain 82 PCM samples instead of the normal 80 of DECT. The remaining 1% will have 84 PCM samples. This may not strictly be true since there is no synchronization between the clock generating the beacons in the AP and the clocks used to take the audio samples. Some difference is to be expected and will require samples to be added or dropped from time to time.

IV. PROTOCOLL ADAPTATION LAYER (PAL)

The PAL interfaces the telephony protocol stack and the WLAN radio while considering QoS support (Fig. 3). The PAL must map the functionality and requirements of DECT into that of 802.11. In some cases, like authentication and encryption, both DECT and WLAN provide duplicate functionality. In these cases the PAL disables one or the other implementation (in WINDECT WLAN encryption is used: WEP – Wired Equivalent Privacy). In other cases DECT and WLAN have similar but slightly different concepts and the PAL simply maps between the concepts. In other cases, like synchronization and handover, the concepts are totally different and the PAL has to hide these differences while still providing the required service.

In the following we give some examples of important functions of the PAL:

Mapping: The PAL has to map between the DECT addressing scheme, used by the higher DECT layers, and the WLAN MAC addresses. In addition the DECT standard defines a number of primitives between the DLC layer and the MAC layer. These DLC primitives must be

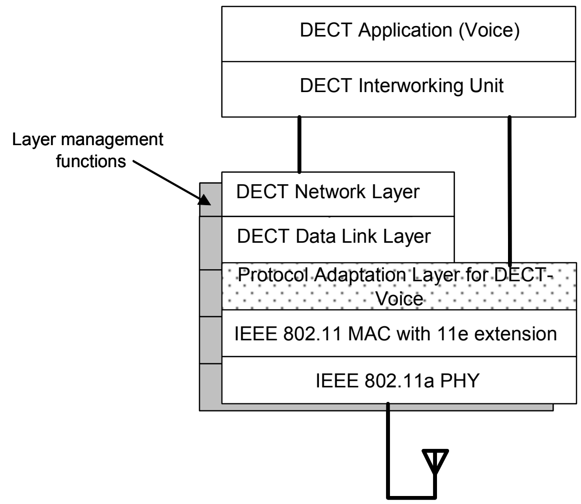

Fig. 3. The WINDECT protocol Stack

mapped into WLAN primitives to enable the DECT stack to continue working, when using a WLAN as the lower layers, The WLAN MLME (MAC sublayer management entity) primitives [1] have no equivalent in DECT, e.g. WLAN defines authentication in the MAC layer, in DECT it is defined in the NWK layer. In such cases the PAL triggers the MLME primitives of the WLAN MAC when such higher level DECT primitives are invoked [10].

The IEEE 802.11 WLAN does not have connections in the same sense as DECT does; the data transmission is based on a frame by frame base. Therefore it is required to use a PAL peer-to-peer protocol to establish, maintain and release connections. The messages of this protocol will be carried in MA-UNITDATA.request/indicate messages [1].

The WINDECT concept introduces some new IEEE 802.11 MLME primitives. Two of these are used to set and report a WINDECT specific information element in beacons. This information element broadcasts paging and cell load information. A similar primitive is used for the neighbour channel information element (these elements are described in the following paragraph about handover).

Seamless Handover: A handover can be performed if an AP has nearly reached its traffic capacity limit, i.e. to aid load balancing, or if a terminal is moving and the signal quality becomes poor. DECT telephony networks support seamless handover. A DECT terminal is able to have connections to two different base stations. This allows a fast handover if the connection to the base station, that is currently used for the communication, gets weak. This is not standardised for IEEE 802.11, where a STA can only be associated to one AP. The WLAN standard supports only break-before-make handover, this can lead to delays that affect speech quality.

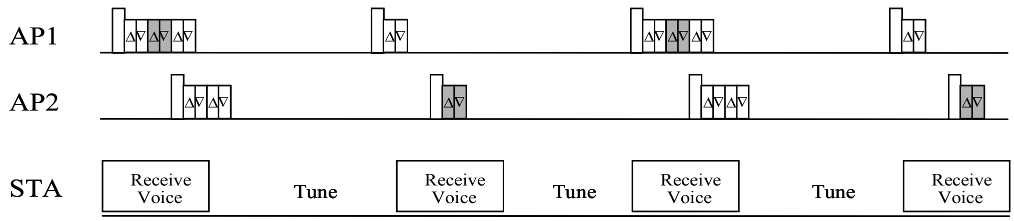

Fig. 4. Seamless handover using active connections to two APs

In order to be able to perform a seamless handover, a terminal has to maintain a list of possible alternative APs. In the WINDECT approach, maintaining this list is achieved by occasionally scanning for other AP between the delivery of voice packets to the current AP and by information about other APs in the vicinity provided by the current AP. The scheme for beacon transmission described in section III guarantees that the beacon from a second AP can be received. The scheme uses a beacon period of 10.24 ms, but voice packets are sent every other beacon. Voice data frames to an active STA are transferred in the CFP, which ends not later than 5.12 ms after the beacon (see section III). If retuning of the receiver to a given frequency lasts about 2 ms, than about 11 ms may be spent on another frequency listening for beacons of other APs. This is sufficient time for receiving the beacon and transferring data. In Fig. 4 this novel approach to seamless handover for WLAN is shown.

During handover, the terminal must be able to maintain an association and a connection to two APs at the same time. This requires an extension to the WLAN standard since such an arrangement is currently not allowed according to IEEE 802.11 [1]. This is sensible as far as pure data transfers are concerned, using the distribution system (DS) [1]. In WINDECT however, the STAs are exchanging their voice traffic with the APs, not with devices on the DS; WINDECT pure voice terminals do not even require the presence of a DS. In WINDECT during the handover process, the DS is not informed about the association to the second AP at least until the handover is completed. After completion the STA has only one association, and the DS is informed that the STA is now served by another AP.

The use of IEEE 802.11f [8] speeds up the authentication process by introducing neighbour graphs, as described in [8]. For WINDECT in addition the (re-) association response message, sent by an AP to a STA, contains information about other APs in the vicinity and about the channel they are using (for more details about the handover process in WINDECT see [10] and [11]).

Load Balancing between APs: An AP has a fixed capacity; in WINDECT up to 20 voice channels can be served by one AP in the CFP. When this capacity is reached, further terminals cannot be supported without decreasing the quality of existing connections. Trying to perform a handover to a fully loaded AP will fail. In WINDECT the AP will broadcast load information at regular intervals so allowing a terminal to know a-priori if the AP has sufficient resources to accept a new terminal (in particular how many additional STAs could be served). The load information may also be used to request terminals to search for another suitable AP. There are situations where one AP may be nearly fully loaded but the neighbouring APs are lightly loaded. Asking a terminal to move to a lightly loaded AP could allow a terminal with no choice of AP to remain connected as it moves from one coverage area to another. This is illustrated in Fig. 5. The middle AP (labelled ‘A’) has four terminals communicating with it, which for this particular AP is its maximum capacity. The lower terminal (labelled ‘5’) is moving upwards. With the current situation, it would not be possible for terminal ‘5’ to perform a handover since the next AP (‘A’) in its path is fully loaded. With this load information STAs know the load situation of each AP. Therefore at first terminal ‘3’ which can access AP ‘A’ and ‘B’ will choose ‘B’ which allows STA ‘5’ to perform a handover to ‘B’.

A second possibility to use the WINDECT load information is that an AP can request, that terminals associate to another AP if possible. The terminal ‘3’ at the right will see that it can perform a hand-over to the AP labelled ‘B’ which has a lower load. This handover relieves the load on AP ‘A’ and thus illustrates how load balancing allows the lower terminal (‘5’) to perform a handover and not have to drop its connection.

Speech Quality: Speech quality as perceived by users should be indistinguishable from that of DECT with its normal physical layer. Packetization delay with WINDECT is to be made as close as possible to the framing of normal DECT which is 10 ms. Therefore there is no possibility of retransmitting packets which are lost or damaged by the effect of transmission errors. Packet loss concealment (PLC) by extrapolating the evolution of a speech waveform from previous packets is an important consideration.

In contrast to many IEEE 802.11 implementations in which the MAC layer discards erroneous packets, in WINDECT, erroneous voice packets as well as correct ones are forwarded to the PAL so that any useful data in erroneous or damaged packets may be used for error concealment. In order to do this a change in MLME element, i.e. MA-UNITDATA.indication has been considered. The ‘RECEPTION_STATUS’ parameter may report the error code “CRCerror” instead of “success” to indicate that an erroneous packet, i.e. a packet for which

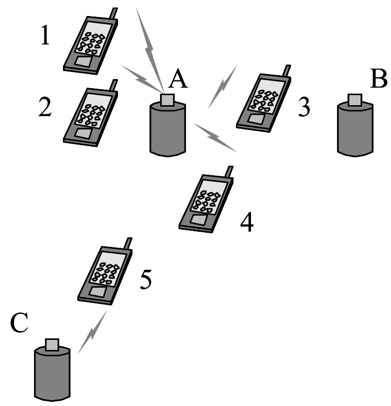

Fig. 5. : Load Balancing between APs

the Viterbi decoder has failed to correct all the bit-errors, is being forwarded to the PAL.

In WINDECT damaged packets will be allowed to reach the higher layers in the 802.11 protocol stack possibly with additional information from the Viterbi decoder. There is no retransmissions (like RTP) and damaged packet delivery will allow novel error correction schemes to be implemented. It is likely that a damaged packet will generally be better than no packet, though the effect of the Viterbi decoder on the bit-stream when it fails to correct all bits will be examined in our future work.

Power consumption: Users of traditional DECT devices have come to expect a long battery life. Talk times of 8 hours and standby times of a week are not unusual. WLAN devices tend to be more power hungry.

Battery life would be poor if a terminal was forced to receive every beacon in order to be able to receive incoming calls. That's why for WINDECT it has been decided not to use the Traffic Indication Map (TIM) [1] to announce a call to a terminal. The use of the TIM would require all terminals to be associated to an AP all the time and to reassociate when the STA moves between different AP's coverage area. From experience with DECT, this is known to be not a good solution since it does not scale to large numbers of terminals. An extra information element will be added to the beacon to indicate a call request from the fixed side and to give load information about the RFP. To avoid the need for a terminal to listen to nearly every beacon and to avoid wasting bandwidth by transmitting the same information in many beacons, a multiplex scheme similar to the one DECT defines, is used. To be able to react relatively quickly on call notifications, WINDECT must use a paging scheme as defined in DECT. DECT paging is based on a 160 ms multi-frame period. The first frame in a multi-frame can contain paging information. Applying the same rule for the WLAN beacon means that a terminal must synchronise to a specific WINDECT-info multiplex in the beacon and has only to listen at intervals of 160 ms for one beacon if the same reaction time as obtained in traditional DECT is to be achieved.

III. CONCLUSIONS

In this paper a new approach to voice over WLAN is given, based on a combination of the upper layers of the DECT standard and the MAC and PHY layer of the WLAN IEEE 802.11. The WINDECT extensions to IEEE 802.11a/e allow that present DECT speech solutions can easily be adapted to the 5 GHz band by exchanging the DECT MAC and PHY with WINDECT PAL and WLAN MAC and PHY. A DECT implementation based on WINDECT can be easily combined with typical IP/ WLAN data applications using the same hardware. Data rate restrictions of the present DECT standard are no longer valid. Several new features of WINDECT have been considered in this paper. In WINDECT handover, speech optimisation, power consumption and load balancing have been employed in a different way compared to other WLAN systems. The latter feature is also quite different from DECT. These characteristics make WINDECT unique among other systems.

IV. ACKNOWLEDGEMENT

The authors would like to thank all partners of the WINDECT project, that is partially funded by the European Commission and Swiss Federal Office for Education and Science, for their contributions and constructive discussions.

REFERENCES

[1] IEEE Std. 802.11, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications", 1999 Edition (ISO/IEC 8802-11:1999).
[2] IEEE Draft Standard 802.11e / D8.0, "Medium Access Control (MAC) Quality of Service (QoS) Enhancements", 2004.
[3] S. Mangold, S. Choi, P. May, O. Klein, G. Hiertz, L. Stibor, "IEEE 802.11e Wireless LAN for quality of service", European Wireless 2002, Florence Italy, Feb. 25-28,2002.
[4] J. A. Phillips and G. Mac Namee, "Personal wireless communication with DECT and PWT", Norwood, MA: Artech House, 1998.
[5] http://www.windect.ethz.ch
[6] IEEE Standard 802.11a, "High-speed physical layer in the 5 GHZ band", 1999.
[7] IEEE Standard 802.11h, "Spectrum and transmit power management extensions in the 5 GHz band in Europe", 2003.
[8] IEEE Standard 802.11f, 2003.
[9] S. Sharafeddine, A. Riedl, J. Glasmann, J. Totzke "On traffic characteristics and bandwidth requirements of voice over IP applications", ISCC 2003.
[10] A. Lunn, K. Tittelbach-Helmrich, F. M. Krause, T. Becker, B. M. G. Cheetham, M. Kuhn, M. Methfessel, A. Ettefagh, M. Secall, M. Spegel, "Design of PAL and voice optimization for WLAN", 2004. Online available: http://www.windect.ethz.ch/documents/
[11] A. Lunn, Michael Kuhn, Marc Kuhn, A. Ettefagh, B. M. G. Cheetham, M. Spegel, "High quality telephony over WLAN", ECWT 2004, Amsterdam, to appear.