Draft Version of an article that appeared in: IEEE Personal Communication Systems in: June 1998

Power Saving Mechanisms in Emerging Standards for Wireless LANs: the MAC Level Perspective

Hagen Woesner, Jean-Pierre Ebert, Morten Schläger, Adam Wolisz*
Technical University Berlin
Telecommunication Networks Group†

April 5, 2001

Abstract

This article provides an overview of mechanisms used for power saving in the upcoming standards for wireless LANs: IEEE 802.11 and ETSI RES 10 HIPERLAN. Power saving on the MAC level is addressed by these standards in a quite different way. We will outline the main features of mechanisms in both standards in terms of power saving. In addition to this we present simulation studies of the power saving mechanism in ad hoc configurations of IEEE 802.11 networks, which demonstrate the optimization potential and some performance trade-offs quantitatively.[1]

1 Introduction

Wireless Local Area Networks (WLANs), with their claim to offer a shared-medium bit rate in the magnitude of typically 1-10 Mbit/s to hosts with mobility limited to in-door or plant environment, belong to a quite new area in wireless communication, with acknowledged high growth potential [1]. Although mature solutions have been commercially available for several years (see [2],[3]), surprisingly enough the expected rapid deployment of WLANs has not taken place so far. Two reasons for this phenomenon are frequently proposed:

Up to now only proprietary solutions are offered, which are not mutually interoperable. For several years owing to the evolving standards IEEE 802.11 [4] and HIPERLAN [5] there have been high expectations concerning the unification of the products. So numerous potential users decided to delay buying until products conforming to these standards would appear. At the moment manufacturers have different levels of commitment to supporting the emerging standards: recently a large number of manufacturers announced the introduction of IEEE 802.11 conforming products for the near future, while only one serious announcement to support the technically much more ambitious (and therefore more expensive in manufacturing) HIPERLAN in products is known to the authors.

The second reason is in fact the power consumption issue. A basic system requirement of WLANs is the support of portable devices like notebooks with significant processing power (typically with a Pentium or PowerPC class processor) or Personal Digital Assistants (PDAs) rather than simple devices like

*also with GMD Fokus, Berlin

†email: {woesner, ebert, morten, wolisz}@ee.tu-berlin.de

[1]This work was partially supported by the Deutsche Forschungsgemeinschaft (DFG) under the following grants: Priority program 'Mobile Communications' and GRADUIERTENKOLLEG 'Kommunikationsbasierte Systeme' (GRK-86) and by the Bundesministerium für Bildung und Forschung (BMBF) under the ATMmobil Project (522-4001-01 BK 615/7)

pagers or mobile phones. From this point in the article we will refer to devices of this class using the term *portables* or *end systems*. Even in stand-alone mode, one of the critical limiting operational factors for portables is their operation time, restricted by the battery capacity. It is generally agreed, that these portables should allow a working time of 4-6 hours without the need to recharge. The computer/communication industry tries to parallel the (notable) effort to increase the longevity of batteries by attempts to reduce the power consumption in all the aspects of portable device design, but in fact only the newest generation of portables approaches this goal with tolerable battery size and weight.

Unfortunately, if modern portable computers are connected to contemporary WLANs the longevity of the battery can be significantly reduced. In [6], it has been pointed out, that contemporary wireless network interface cards (with 1 Mbit/s bit rate) can take 12 times more power than a standard 10 Mbit/s ethernet card. Thus it should not be surprising that longevity reduction in the range of 60% has been reported for portables in [7]. Depending on the method of usage, we observed during our lab work a dramatic drop of time in action from 3 hours to 45 minutes with a laptop using a WLAN PCMCIA interface card.

In general the issue of power saving (PS) in mobile communication is subject to numerous activities (see e.g. [8],[9],[10],[11],[12],[13]), so what is special with WLANs? In agreement with the widely accepted IEEE 802.x approach a LAN can be logically divided into the Physical (PHY), the Media Access Control (MAC) and the Logical Link Layer (LLC). Physical Layer solutions for WLAN are in fact not so different from those used in other wireless communication systems. On the other hand what makes the WLANs so different from other mobile communication systems are link layer solutions, especially MAC. And it has been recently recognized (e.g. [7],[11]) that protocols might significantly influence power consumption of mobile systems.

The aim of this article is to discuss how the problem of power saving has been addressed in the IEEE 802.11 and HIPERLAN standard drafts, with special regards to the MAC layer solutions. The remainder of this article is organized in the following way:

In the next section we present an overview of main power saving issues on the MAC level. Since power saving on MAC level is not independent from the power consumption aspects of underlying PHY layer services, some of these aspects will be mentioned, too. A short description of the IEEE 802.11 architecture and a fairly detailed discussion on how the power saving problem is addressed in the MAC of this standard follows. Further we present a simulation study demonstrating influence of individual parameters of the power saving mechanism on the system Quality of Service for the basic variant of the IEEE 802.11 approach (which from the power saving point of view is surprisingly the most complex!)

After that we present the architecture and a detailed description of how the power saving problem is addressed in the MAC protocol of HIPERLAN. We also focus on the differences to IEEE 802.11 and their possible implications. The article is completed by a discussion of open issues, especially in the context of harmonization of the power saving approaches within MAC with the operation of higher protocol layers typical for a LAN communication software.

2 MAC power saving issues

Roughly speaking, the MAC Layer is responsible for data framing, efficient sharing of the channel and possibly some error control. We identify three basic classes of approaches to power saving in MAC layer:

- Optimized use of the PHY layer services,
- Optimized media access protocol structure,
- Optimized system design.

2.1 Radio physical layer implications

In general, different PHY layer services have different power requirements. Circuitry implementing this function is located on the network interface card (NIC) powered from the portable, and can consume an important amount of energy. So proper use of this services by MAC is critical.

Higher power consumption is in general required for

higher bit rates: one of the reasons is the frequently required higher equalization complexity needed to deal with the intersymbol interference at higher bit rates [14].
Further, Stemm et al. showed in [7], that in a typical useage scenario by far the most power (about 90%) is drawn by listening to the radio channel! Sending or receiving of packets adds only a few percent to the energy balance.

2.2 MAC protocol design

There are several options for an energy efficient MAC protocol design. In general simple protocols need relatively less power than complex protocols. For example, a large number of necessary control messages negatively influence energy per "useful" bit relation. Also, more processing work has to be done. In the following we outline some MAC design options, which in our opinion have the most impact.

2.2.1 Packet structure

Due to excessive long headers (addresses, control fields etc.) or trailers (checksums) the energy per "useful" bit relation may be negatively influenced. One solution to this problem is header compression as it is used for several protocols. Another idea regarding the packet structure is to split the packet into a *low bit rate part* for control information (e.g. addresses) and a *high bit rate part* for data. The intention is to invoke costly functions only when they are needed. For example, the MAC evaluates first the low bit rate control information like the destination address. Receiving of data, which might be costly due to high bit rates (e.g. necessity of equalizer), is only performed if the packet is intended for the end terminal.

2.2.2 Awake/doze mode

In order to save energy the network interface card (NIC) may be switched off when there are no transmissions (doze mode). Otherwise the NIC is in the awake mode. Following this idea, the use of the doze mode has the potential to improve the power save gain substantially. Unfortunately, one can not do this without losing the capability to communicate in both directions, i.e. a station in this kind of a power saving mode would not know of any data arriving for it during this time. Nevertheless, it is possible to switch off the NIC if two problems are correctly addressed in the MAC protocol design:

- How does a station make sure to receive packets from other stations, even if it is in sleep mode most of the time?
- How does a station send data to another station that is in sleep mode?

Beside the power save impacts on the MAC design there are impacts (e.g. timer) on other protocol layers.

2.2.3 MAC level error control

The channel quality is often improved by forward error correction mechanism (FEC) on the PHY layer. If the offered channel quality is still not satisfactory, MAC level retransmissions may be used. Since the radio channel quality may be persistent for a while (good or impaired), retransmission of MAC packets in the impaired radio channel state is unnecessary and therefore expensive. Transmission channel probing, first proposed by Zorzi in [15], can be used to overcome this problem. The idea is simple: Instead of retransmitting the MAC packets again over an impaired radio channel, short low power probe packets are sent continuously unless feedback is received for these probe packets. When one probe packet has been successfully transmitted indicating an appropriate channel quality, the retransmission of the data packet is scheduled. In other words, retransmission of data on a relatively high power level is only scheduled if the channel quality is sufficient. As a result, no energy is wasted for retransmission during channel impairments.

2.3 System design

The critical system design issues for power saving with respect to MAC are:

2.3.1 Radio cell diameter

The radio cell diameter is defined by macro (several km), micro (up to 1 km) and pico (a few meters). In general the smaller the radius is, the less is the power usage. This is obvious, but also shown in [6], where five different WLAN products using similar technology are compared. The difference with regard to power drain is the signal strength and therefore the range. Because of lower power consumption, higher possible bit rates, and a better frequency reuse, WLANs tend to be made up of small cells. However the efficient use of the pico-cells is critically dependent on an efficient extension of the coverage beyond the border of a single radio cell. The coverage issue in IEEE 802.11 is addressed by a distribution system (not specified in the draft standard). HIPERLAN addresses the coverage issue with a forwarding method, which is explained in more detail in one of the following sections.

2.3.2 System architecture

WLANs can be classified as distributed or centralized systems, which are also referred to as ad hoc and infrastructure based systems. Centralized systems, consisting of a base station and several portables, are inherently more suitable for the design of low power consuming end systems [17]. The reason for that is the base station, which can be equipped with more intelligence and sophisticated (perhaps significantly more power consuming) hardware since it is normally fixed to a certain place. Therefore power supply is not the problem. In doing so, portables can be off-loaded in terms of MAC functionality and power hungry processing hardware. This is referred to as asymmetric design [18],[19]. The drawback of this method is the more limited flexibility in contrast to distributed systems. Furthermore, MAC PDUs are usually transmitted via the central control unit. The disadvantage of this design arises whenever two portables in the same radio cell communicate with each other (e.g. portable→base station→portable). In this case data has to be transmitted twice instead of once (portable→portable), which leads to a waste of bandwidth and energy as well as to an increased risk of data corruption. Neither IEEE 802.11 in the basic operation mode nor HIPERLAN use the asymmetric design option for MAC operation. However, the optional Point Coordination Function of IEEE 802.11 applies this design to a certain degree. As we will see later, low energy consumption is often compromised by the performance of the MAC protocol. Considerations of power saving issues are fundamentally influenced by the trade-off between the energy consumption and achievable Quality of Service (throughput, delay, error rate) for predefined distance, coverage and bit rate. The aim is to transmit with as little energy as possible while meeting the required Quality of Service. There are several measures for energy consumption. A portable's energy efficiency measure can be the number of delivered useful data divided by the consumed energy or the consumed energy per time [15]. This strongly depends on measurement assumptions and system parameters. To show the energy consumption of a portable, other abstract measures (e.g. time in sleep mode vs. delay, throughput, etc.) may be used.

3 Power Saving in the IEEE 802.11 draft standard

3.1 Overview

Work on the IEEE standard for Wireless Local Area Networks started in 1990. The draft standard covers the lower two OSI-Layers. It defines a common medium access protocol for three different physical layers: Infrared (IR), Frequency Hopping (FHSS) and Direct Sequence Spread Spectrum (DSSS), each capable of a data rate of 1, optional 2 Mbit/s. There are two modes of operation depending on the existence of an access point (AP) in the BSS (Basic Service Set - a wireless cell). They are referred to as the Distributed Coordination Function (DCF) and the Point Coordination Function (PCF). The DCF is the basic access algorithm, it has to be supported by all stations in the network. The PCF is optional and supports time-bounded services such as audio traffic.

It builds a contention free period on top of the basic access mechanism, which is contention based. A Distribution System (DS) working on a logically separated medium connects the access points of the BSS to form an Extended Service Set (ESS). A framework for a distribution system is provided in [20] and one possible solution for it is shown in [16].

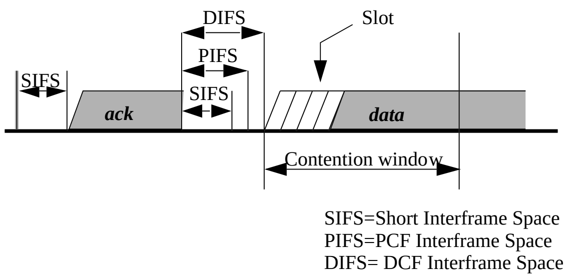

Figure 1: Medium Access in the Distributed Coordination Function

Medium Access in both, the DCF and the PCF, is based on a general MAC, which was approved by the working group in 1994. It is called DFWMAC and belongs to the class of CSMA/CA protocols (Carrier Sense Multiple Access/Collision Avoidance). Since this algorithm is fundamental to IEEE 802.11, it will be explained in short at this point: A station X that has to transmit a packet selects a slot n out of a certain (medium dependent) number of slots (equally distributed) and monitors the channel until the end of the (n-1)th slot. If no other station has started transmitting on the channel up until this time, station X starts its transmission in the nth slot. If another station had started transmitting earlier, which means that it had selected a smaller slot number, transmission is deferred until the end of that packet exchange. The slot number in the next transmission attempt is decremented by the number of idle slots in the previous cycle. This gives higher access priority to stations that have been in the competition for a longer time.
Information about the state of the channel is available from two sources: the actual carrier sensing and the so-called Net Allocation Vector (NAV) which is a virtual carrier sense mechanism. The NAV is local to each station and indicates a busy channel for future traffic. There is an optional exchange of RTS/CTS (Request/Clear To Send) packets prior to the transmission of the data packets. This is useful in the hidden terminal case, where all stations update their NAV with the information about the duration of the next transmission in either the RTS or the CTS packet. That way silence around the receiver can be guaranteed even if not all stations in this area received the RTS packet correctly.
Priorities in the access to the medium are translated into interframe spaces. In the PCF, the AP has to have a higher access priority than all other stations. It therefore waits for a shorter time before deciding that the channel is free (see the PCF-Interframe Space in Figure 1). This way it can set up a superframe structure to support time-bounded traffic. The first part of the superframe is reserved for time bounded traffic. The AP sends downlink packets and polls stations for uplink packets. In the second part of the AP-created superframe medium access is performed using the DCF (see Figure 1). A more detailed investigation of the access and the RTS/CTS mechanism can be found in [21], [22] and [23].

3.2 Timing Synchronization and Power Saving

Within the standard, the general idea is for all stations in Power Save (PS) mode to switch off the radio part for some period. They have to be synchronized to wake up at the same time when there starts a window in which the sender announces buffered frames for the receiver. A station that received such an announcement frame stays awake until the frame is delivered. This is easy to be done in the PCF, where there is a central access point which is able to store the packets for stations in doze state and to synchronize all mobile stations. It is much more difficult for the DCF, where the packet store and forward and the timing synchronization has to be done in a distributed manner.
Power Saving in IEEE 802.11 therefore consists of a Timing Synchronization Function (TSF) and the

actual power saving mechanism. The TSF for an infrastructure network (the PCF) can be seen in Figure 2. The access point (AP) is responsible for generating beacons which along with other information contain a valid time stamp. Stations within the BSS adjust their local timers to that time stamp. If the channel is in use after the beacon interval the AP has to defer its beacon transmission until the channel is free again. The power management in the PCF is simple due to the existence of the AP as a central buffer for all packets destined for the stations in doze mode. Along with the beacon the AP transmits a so-called Traffic Indication Map (TIM). All unicast packets for stations in doze mode are announced in the TIM. Afterwards the mobiles request the packets from the AP. If broadcast/multicast frames are to be transmitted, they are announced by a Delivery TIM (DTIM) and are sent immediately after. Of course the stations in power save mode have to wake up shortly before the end of the beacon interval and to stay awake at least until the beacon transmission is over.

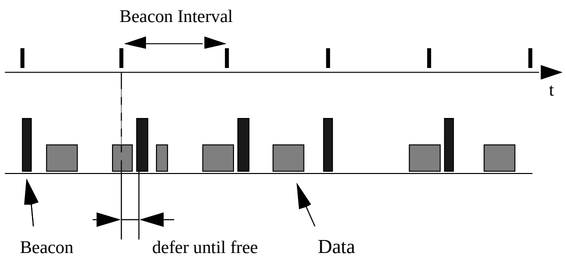

Figure 2: TSF for infrastructure networks in 802.11

The TSF is more complicated for an ad hoc network (the Distributed Coordination Function - DCF, see Figure 3). Due to the absence of a trusted authority the timers adjust in a distributed way: Every station is responsible for generating a beacon. After the beacon interval all stations compete for transmission of the beacon using the standard backoff algorithm. The first station "wins" the competition and all others have to cancel their beacon transmission and to adjust their local timers to the time stamp of the winning beacon.

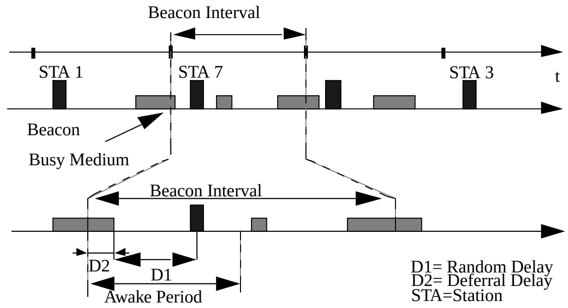

Figure 3: TSF for ad hoc networks in 802.11

The power management in the DCF is based on the same distributed fashion as used for the TSF. Packets for a station in doze state have to be buffered by the sender until the end of the beacon interval. They have to be announced using Ad hoc TIMs (ATIMs), which are transmitted in a special interval (the ATIM window) directly after the beacon. ATIMs are unicast frames which have to be acknowledged by the receiver. After sending the acknowledgment, the receiver does not fall back into doze state but stays awake and waits for the announced packet (see Figure 4). Both ATIMs and the data packets have to be transmitted using the standard access algorithm.

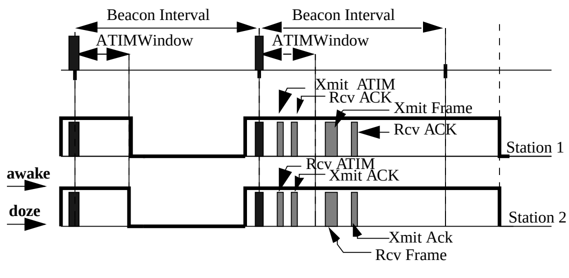

Figure 4: Power Management in the DCF of 802.11

4 Simulations of the Power Saving Mechanism of the IEEE draft 802.11

4.1 Simulation Approach

Since the DCF is the basic access algorithm we decided to perform simulations of this most important (and in addition more complex) scenario.
Our simulations were performed by using the PTOLEMY simulation environment [24]. We used the appropriate values for the Direct Sequence Spread Spectrum (DSSS) physical layer.
The simulation environment consists of 8 stations, which belong to an Independent Basic Service Set (IBSS). The stations are situated within a diameter of 100 m. We simulated exponentially distributed packet losses with a probability of 10^{-5}. We did not consider any hidden terminals. Simulations with 1, 2, 4 and all 8 stations in power save mode were performed.
Since one of the areas of deployment of wireless LANs will be the replacement of traditional copper wiring, we used as sources trace files of an Ethernet traffic. The trace files were recorded on our institute-internal 10Base2 Ethernet. We scaled the traces to simulate overall offered loads of around 15, 30 and 60% of the 2 Mbit/s raw physical throughput.
Our aim was to tune the algorithm by varying the beacon interval and ATIM window size (ATIM window<beacon interval) to investigate the trade-off between the throughput of stations in power save mode and a maximum possible time in doze state. We chose the ratio of time in doze state versus the time in active state as a measure for the quality of the power saving mechanism itself.
Any additional effects which are depending on the PHY layer, such as equalization and on-off switching costs, could not be taken into account, because of the diverse nature of the physical layers.

4.2 Simulation Results

First we wanted to observe dependencies of different sizes for the beacon interval and of the ATIM window on the throughput. As it may be expected, higher numbers of stations in the power save mode lead to lower throughput. This is because of the overhead for each data packet transmission, which consists of an ATIM and an ACK and two backoff sequences, regardless of the size of the packet to be transmitted. It showed that there is a decrease in throughput for very small and very large ATIM window sizes (see Figure 5). An ATIM window which is too small results in less ATIMs and therefore in less packets, which can be announced and transmitted. On the other hand, when the ATIM window is too large, more ATIMs are sent than there is actually time for the packets. When we used a lower offered load for the simulations the results were basically the same, though throughput was constant for a broader range of ATIM window sizes. This was due to the fact that the channel could not be saturated any more.

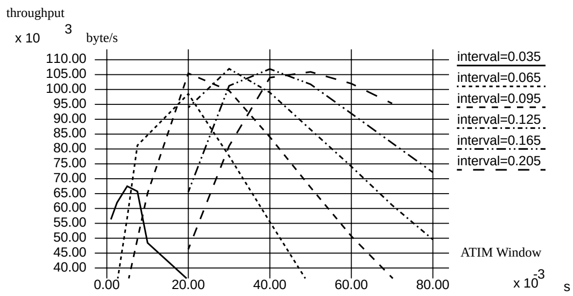

Figure 5: Throughput vs. ATIM window size for different beacon intervals, load=60.76%, 8 stations in power save mode

All our simulations led us to the "rule of thumb" that the ATIM window size should be proportional to the beacon interval and that it should take approximately 1/4 of the beacon interval.
The next question was to determine the time in doze state in relation to the total time. In Figure 6 one can see that the time in doze state increases when using shorter beacon intervals. The simulation shown here was performed at an offered load of about 30%.

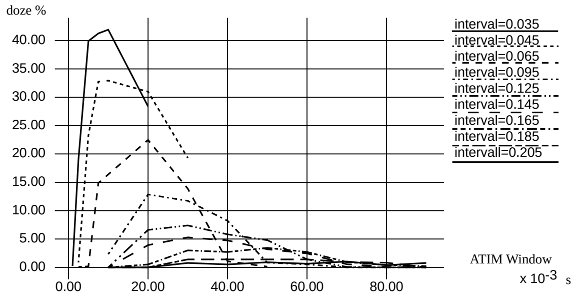

Figure 6: Percentage of time in doze state vs. ATIM window size for different beacon intervals, load=30.72%, 8 stations in PS mode

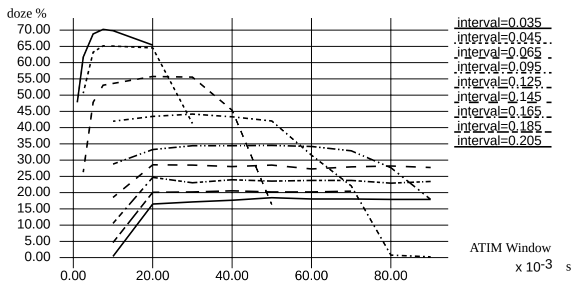

Figure 7: Percentage of time in doze state vs. ATIM window size for different beacon intervals, load =15%

In Figure 7 we present simulation results for the same scenario as before, but with an offered load of about 15%. It shows that a station can stay in doze mode up to 70% of the time for beacon intervals small enough to allow for a fast transmission of the packet. The results can be explained as follows: The bigger the beacon interval the bigger the possibility that a station wishes to send during that time. This means that it has to transmit ATIMs in almost every beacon interval and to stay awake until the transmission is completed. The same applies for a receiving station. In addition to that, more ATIMs per beacon interval have to be transmitted in bigger beacon intervals, which leads to a higher collision rate and longer medium access time.

4.3 Discussion of the Simulation Results

We presented simulations of the power saving mechanism in ad hoc networks using the IEEE 802.11 standard. Work on this simulations started at a time when there was no recommendation for standard values for beacon interval and ATIM window size in the then current version of the draft standard. In the meantime these values are set to be 100 ms for the beacon interval and only 4 ms for the ATIM window. Based on this work we can recommend figures for the ATIM window and beacon interval. Generally the mechanism gets less sensitive against the ATIM window size with higher values for the beacon interval. The simulations showed an optimum for the throughput at about 95 ms beacon interval. This figure corresponds well to the results of optimum handover beaconing shown in [25]. The beacon interval should be smaller to lead to longer times in doze state. There is a trade-off between power saving and the overhead needed for it. If we were to sacrifice about 10% in throughput we could save up to 30% energy. Smaller beacon intervals, however mean a more frequent switching between the doze and awake state. In line with the switching cost between the two states, the optimum for the beacon interval again increases.

The ratio between ATIM window and beacon interval should be around 1/4. While the first result corresponds to the value in the draft quite well, there is a remarkable difference in our recommended size for the ATIM window. This should be explained as follows: The recommended ATIM window size of only 4 ms (or Kμs, according to the standard) will be too small if there are a number of stations in power save mode or if the overall load is above 10%. In our specific scenario we would definitely recommend a higher value of the ATIM window parameter.

There should be a means to adapt the the ATIM window size to the offered load or, to be more exact, to the sum of the offered loads of the stations in power saving mode.

Although the draft standard states that the ATIM interval should be static during the lifetime of an IBSS, an initial idea for adjusting the ATIM window size dynamically could be the following: The offered load of the stations in power save mode corresponds to the number of the winning slot in the contention window. If many stations generate (not necessarily send) an ATIM packet, then there is a high probability for a low slot number winning. Therefore the offered load could be estimated by the mean of the winning slot number. Every station would have to calculate the new ATIM window size out of this mean. Because ATIM window size is part of the IBSS parameter set transmitted within the beacon, a distributed and dynamic adjustment of the ATIM window size should be possible.

At the time of writing this article there was some discussion within the 802.11 working group about the usefulness of this power saving mechanism. The basic argument was that the sending of an ATIM packet could in principle be delayed indefinitely due to the CSMA/CA mechanism.

Although this is theoretically true, it would apply to any packet being transmitted using a CSMA-like technique. Our simulations show that power saving in the Distributed Coordination Function is feasible and useful, though there might be some argument about certain values of parameters.

5 Power Saving in HIPERLAN

5.1 Overview

HIPERLAN is the WLAN specified by ETSI RES 10 [5]. Like the IEEE 802.11 standard the HIPERLAN standard defines medium access control and the physical layer. However, the design goals behind the two standards are different. The basic idea behind HIPERLAN was to develop a wireless LAN which can operate completely independent from any infrastructure while supporting both ad hoc networking and complex networks, which are composed of multiple cells, without distinguishing between two different modes (namely ad hoc and infrastructure based mode) like IEEE 802.11 does. Thus, HIPERLAN does not need a central station (base station) in order to allow range extension or to support different service classes.

The frequency band used by HIPERLAN is in the 5.2 GHz area and is divided into five independent channels. The physical layer operates at two different data rates - a low data rate (1.4706 Mbit/s) used to transmit acknowledgment packets and the packet header and a high data rate (23.5294 Mbit/s) to transmit the data packet itself. The reason for using two different data rates is explained below in the power saving chapter.

The data transfer function of the MAC supports both asynchronous and time bounded data transmissions. This is achieved by specifying a priority for each data packet by means of the data's lifetime and a flag indicating low or high priority. The lifetime of the data describes the time by which the packet must be delivered to the receiver in order to be of any use for the receiver. The medium access control of HIPERLAN takes care that packets with a shorter residual lifetime are transmitted first. To achieve this the residual lifetime and user priority are mapped to a channel access priority ranging from zero (highest priority, residual lifetime less than 10 ms) to four (lowest priority, residual lifetime more than 80 ms). All packets with a residual lifetime equal to zero are discarded either by the source or by a forwarder.

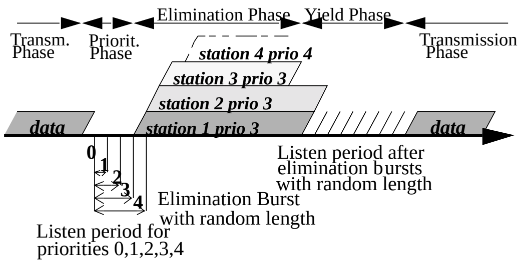

Figure 8: EY-NPMA access mechanism

Like IEEE 802.11 HIPERLAN uses a kind of CSMA/CA to regulate the channel access. The variant used by HIPERLAN is called EY-NPMA (Elimination Yield - Non-preemptive Priority Multiple Access). EY-NPMA splits the access procedure into three phases (see Figure 8): Priority Resolution, Elimination and Yield Phase. In the Priority Resolution phase a station listens to the medium for the Priority Detection Period which has a length proportional to its priority. If a station detects a signal during this time it stands back from transmission. Otherwise it transmits a burst (a well defined bit stream, which is transmitted by using the high bit rate) to signal its priority to other stations. All stations with the same (highest) priority survive this phase. In the Elimination Phase each surviving station sends an elimination burst the length of which is bound and defined by a certain discrete probability distribution. After sending its burst a station switches to listening mode. If a station still detects a burst on the medium after sending its own one it defers from the access cycle, otherwise it survives this phase. In the Yield Phase every surviving station listens to the channel. Again the listening period is individual for every station, bounded and defined by a certain discrete probability distribution. If a station hears a signal during this period it stands back from transmission, otherwise it transmits its data frame immediately after the yield period. A more detailed description and some performance results can be found in [26].

Beside the channel access mechanism one should notice some other special features of the HIPERLAN MAC in order to understand some of the problems regarding power saving discussed in the next chapter.

First of all due to the renunciation of an infrastructure a mechanism is needed which allows the physical extension of a HIPERLAN beyond the radio range of a single station. Therefore the HIPERLAN standard defines a forwarding mechanism. This mechanism does not need to be implemented in every HIPERLAN node. Stations that support forwarding are called forwarders. The information needed for forwarding is maintained by use of the Routing Information Exchange functions of the MAC.

A logical consequence of the forwarding concept is that the physical size of a HIPERLAN is defined by the positions of the forwarders and non-forwarders. Due to the mobility of every station (including the forwarders) the physical size of a HIPERLAN is a function of the current position of all stations.

While forwarding solves the problem of the limited radio range of a single station it introduces some new problems. An additional amount of data has to be exchanged between the nodes in order to keep track of the network topology. Further we need overlapping cells to enable forwarding which leads to the hidden terminal scenario [26]. Finally with regard to power saving, it is obvious that the power consumption of a forwarder increases, because it has to receive, buffer and forward packets which are sent to one of its clients. This raises the question: "Why should any node support any other node and therefore increase its own power consumption?" It is obvious that forwarders can not be battery driven but have to be connected to a power supply. Thus, forwarders are not expected to be mobile. Another difficulty due to the mobility and limited radio range is that a fragmentation situation may occur in which a HIPERLAN is effectively partitioned into multiple disjoint communication subnets. Therefore a mechanism is needed to automatically re-merge fragments of a HIPERLAN when possible. On the other hand it is also possible for multiple different HIPERLANs to use the same radio channel. In such a situation it must be possible to distinguish the traffic of the different HIPERLANs. In order to overcome these two problems each HIPERLAN has a unique identifier. This identifier is also used to join a specific HIPERLAN by the means of the HIPERLAN Lookup functionality in the MAC.

5.2 Power saving

In our introduction we pointed out that a device has to be turned off in order to save power. In the previous chapters we explained how IEEE 802.11 deals with this difficulty. In this chapter we will give an overview how HIPERLAN faces this problem. We just provide an overview since, as already mentioned in the introduction chapter, it seems that HIPERLAN with respect to available WLAN products plays a less important role than IEEE 802.11. In this section we will concentrate on differences between the power saving mechanisms of both standards. One of the differences between the two wireless LANs is the fact that HIPERLAN offers a bit rate of over 20 Mbit/s compared to 2 Mbit/s of IEEE 802.11. This bit rate requires equalization. As pointed out in the introduction an equalizer is one of the most expensive parts of the receiver in terms of power consumption. In order to reduce power consumption without the loss of functionality (e.g. reachability) let us consider how the receiving procedure would operate in general: HIPERLAN is based on a broadcast channel. Thus, each station hears all packets which were transmitted within its radio range. Every time the physical layer detects a signal on the medium it has to turn on the power consuming equalizer to receive the packet. After receiving the whole packet the MAC discards the packet if it is not the receiver.This means that energy was wasted due to the unnecessary use of the high speed receiver including the equalizer.

Thus in order to save power the equalizer should only be turned on when the station is the receiver of a packet. To decide whether a station is the destination of a packet without starting the equalizer each packet is divided into a low-bit-rate and a high-bit-rate part. The low-bit-rate part of a packet consists of 34 bits and is transmitted at 1.4706 Mbits/s which does not require equalization.

One of its fields contains an 8 bit hash sum computed over the destination address of a packet. As soon as a receiving station has received this hash sum it can determine whether it is not the receiver. Of course, the computed hash sum is not definite but it guarantees that if it evaluates to false (station is not the receiver) that this result is correct. Only stations that determine that they could be the receiver turn on their equalizer to receive the high bit rate part of the packet. To summarize HIPERLAN overcomes the problem of the power hungry equalizer by using a clever framing scheme.

Next we will consider an additional power saving mechanism of HIPERLAN which is from its basic structure similar to the distributed mode mechanism of DFWMAC. However, there are some important differences which will be pointed out throughout the description.

The overall design of the mechanism is in line with the distributed concept of HIPERLAN. This means that HIPERLAN does not use a single power save server which is part of any infrastructure like a base station in DFWMAC. Power saving within HIPERLAN is based on a contract between at least two stations. The station that wants to save power is called the p-saver and the station that supports this is called the p-supporter. The p-saver is only active during prearranged intervals while the p-supporter has to queue all packets destined for one of its p-savers and schedule the transmissions of these packets during the active intervals of the p-savers.

The situation here is similar to the forwarding mechanism. Again we have some stations (p-supporters) that have to support other stations whereby they increase their own power consumption. Therefore it is obvious that p-supporters should draw power from a power supply, too. Because of this similarity it could also be expected that every forwarder is a p-supporter as well.

Each p-saver can have multiple p-supporters. All of its p-supporters must be within radio range. The p-saver does not know which stations of its HIPERLAN are p-supporters. Instead of directly addressing the p-supporters, the p-saver broadcasts its request to all neighbors. Thus, it is possible that a p-saver receives the same packet several times from different p-supporters. The duplicates are detected by the use of a HIPERLAN MAC-entity (HM-entity) sequence number. The reason for allowing several p-supporters for each p-saver is to keep the protocol as simple as possible, especially in the case of mobility. The costs for this approach is the waste of bandwidth due to duplicates.

Timing elements	Valid range of value [ms]
Active interval	500 – 65535
Offset	0 – 65535
Period	500 – 65535

Table 1: Valid values

The p-saver informs its p-supporters about its active interval by periodically transmitting a special HIPERLAN MAC PDU (INDIVIDUAL-ATTENTION HMPDU - IP-HMPDU). This PDU contains the length of the interval during which the p-saver is able to receive packets, the amount of time which has elapsed since the most recent start of the active interval (offset) and the amount of time between the start of two successive active intervals (period). The range of values allowed for these parameters is listed in Table 1.

The p-saver transmits IP-HMPDUs periodically for two reasons. First it can update its active interval easily and needs no extra packet to cancel its power saving request. Second, a moving p-saver does not need to know when it leaves the radio range of one of its p-supporters. Further it automatically informs all p-supporters that are new within its radio range about its power saving mode.

The operation of the p-supporter is quite similar to the operation of forwarders. A p-supporter has to receive and store all packets addressed to one of its p-savers. It learns which stations to support by recording the information contained in the IP-HMPDUs from all neighboring p-savers. Finally it must schedule the transmission of the stored unicast packets in line with the active interval of the receiving p-saver. The p-supporter holds only one active interval for each p-saver. Each time it receives a new IP-HMPDU it updates the old information.

The functionality described so far handles unicast packets but is not suitable for multicast and broadcast packets - because of each p-savers' individual active interval the p-supporter would have to send an extra copy for each p-saver. In order to avoid this waste of bandwidth, each multicast packet is only transmitted once. To synchronize the transmission of a multicast packet with all p-savers the p-supporter defines a group-attendance pattern. This pattern is structured like the Individual Attention Pattern of the p-savers and is transmitted regularly to all neighboring p-savers by a p-supporter. Again, the pattern is transmitted periodically because this is the easiest way to keep the state of all p-savers within radio range up to date. The p-supporter has to transfer all its multicast PDUs during its declared recurring active interval. Each p-saver is advised to schedule its reception of multicast packets to this interval. Both p-saver and p-supporter discard the recorded information about individual-attention and group-attendance, respectively, unless they receive an update during a specified time. This can happen for example when a p-saver turns off its power saving mode or when it leaves the radio range of a p-supporter.

To compare this mechanism with the mechanism of DFWMAC one finds that in DFWMAC all stations in power save mode wake up for the same time interval and that a single station is not allowed to define its own individual power saving interval. This means that the queued traffic from all stations is concentrated within the common active interval. Contrary to this, HIPERLAN allows each p-saver to define its own active period and the time between two active periods. Thus the whole network traffic is not concentrated on well-defined intervals. Only for multicast traffic the p-supporters define a common interval in which all p-savers have to be active. On the other hand, unlike DFWMAC, HIPERLAN does not define an announcement frame by which the p-supporter could inform the p-saver about how many packets it has stored. Thus, on the one hand the p-supporter can only use the fixed active interval for transmission which could be too short leading to long delays (the p-supporter has to delay the packets until the next active period of the p-saver) or packet loss due to limited queues. And on the other hand the p-saver has to stay in listening mode during the whole active interval because it has no idea about how many packets it will receive.

Finally we noticed that power saving does not make sense in conjunction with time bounded services. This is due to the fact that the smallest period between two active intervals is 500ms (see Table 1) which is too long for most time bounded traffic.

6 Conclusions

In this article we discussed different aspects of power saving in wireless LANs with special attention paid to the emerging wireless LAN standards, IEEE 802.11 and HIPERLAN. We investigated the mechanisms used for power saving in the two draft standards. As the chances for products to hit the market in the near future are much better for IEEE 802.11, we presented some simulation results for the power saving mechanism in the Distributed Coordination Function of this draft standard. Under fairly realistic assumptions we obtained numerical results which showed a significant difference to the value for the ATIM window given in the standard. We discussed power save options on other protocol layers. Although we basically considered the MAC perspective in this article, it is important to use power saving techniques on other protocol layers. Finally, we outline three relevant approaches in the following:

6.0.1 Inter-protocol adjustment

In [28] it is shown that, under certain packet loss conditions, competing retransmission strategies between link and transport layer protocols lead to a degraded throughput but a higher link utilization. This can be avoided by adjusting the retransmission timers. As a consequence, no energy is wasted for unnecessary packet retransmission.
In [13] the dependencies between forward error correction, typically located in the physical layer, and ARQ mechanisms in terms of energy consumption are shown. ARQ mechanisms are needed to provide for a reliable link, resulting in higher delays and some protocol overhead. FEC is used to increase the link performance and adds some delay, computational cost and overhead in the form of additional bits. There is a trade-off between these two types of error correction depending on traffic type and packet error pattern. Through an adjustment of mechanisms an optimum power save gain can be achieved.
The design of the MAC protocol and of the other protocol layers should therefore be harmonized. These issues are part of our current investigations.

6.0.2 Asymmetric LLC and transport protocols

The idea of asymmetric protocols was first introduced by Ayanoglu et al. [19] with the AIRMAIL protocol. AIRMAIL is a LLC protocol which is based on a centralized system. The dedicated central unit (e.g. base station) holds the larger part of LLC functionality which results in different sizes of object files for the LLC functionality of portables (40%) and central control unit (60%).
Haas applied this idea to the transport layer in [18]. In particular, he developed Mobile TCP (M-TCP), which is an asymmetrical TCP.

6.0.3 Application level

Another method which can be used to improve the power consumption budget is data partitioning [27]. Data partitioning means that data (e.g. on a data base server) is partitioned into broadcast (multicast) and unicast data according to a certain number of requests for that datum (e.g. a data base entry). If a certain datum is requested very frequently, then it is assumed that different portables are interested in that datum. As a result that datum is marked as a broadcast (multicast) datum. It will be transmitted only once in a certain time frame and interested portables do not have to request this datum separately.
This requires an adequate support of broadcast and can result in a substantial power save gain. Clearly the delay for the broadcast data increases which may require delay-tolerant and in the end “wireless aware” applications.

7 Acknowledgments

This article is built upon simulation results which were presented at the ”6th WINLAB Workshop on Third Generation Wireless Networks” at Rutgers University, NJ, March 1997. We would like to thank Christian Röhl for contributing a great share of the simulation work. The working group of Aura Ganz at the University of Massachusetts Amherst helped a lot in a constructive discussion during a preliminary presentation of the simulation approach and results. Special thanks to the reviewers and the editor for their fruitful and valuable comments.

References

[1] K. Chen: "Medium Access Control of Wireless LANs for Mobile Computing", IEEE Network, vol. 8, no. 5, Sept/Oct, 1994, pp 50-63

[2] C.J. Mathias: "New LAN Gear Snaps - Unseen Desktop Chains", Data Communication, March 21, 1994

[3] Wireless For The Corporate User, A Wireless Publishing Co. Publication, Vol. 6, No.1, Jan. 1997

[4] The editors of IEEE: Wireless LAN Media Access Control (MAC) and Physical Layer (PHY) Specification, IEEE 802.11 Draft Version 4.0; May 1996

[5] ETSI, "High Performance Radio Local Area Network (HIPERLAN)", Draft Standard ETS 300 652, March 96

[6] David Newman, Kevin Tolly: "Wireless LANs: How Far? How Fast?", Data Communications, March 21, 1995

[7] M. Stemm, P. Gauthier, D. Harada, R.H. Katz: "Measuring and Reducing Energy Consumption of Network Interfaces in Hand-Held Devices", Proc. of Mobile Multimedia Communication (MoMuC-3), Princeton, NJ, USA, Sept. 1996

[8] A. Salkintzis, C. Chamzas, C. Koukourlis: "An Energy Saving Protocol for Mobile Data Networks", Intl. Conf. on Advances in Comm. and Control, COMCON 5 , June 26-30, 1995

[9] T. Imielinsky , S. Viswanathan, B.R. Badrinath: "Energy efficient indexing on air", Proc. of Intl. Conf. on Managm. of Data - ACm-SIGMOD, pp. 215-26, May 1994

[10] I. Chlamtac, C. Petrioli, J. Redi: "Analysis of Energy-Conserving Protocols for Wireless Identification Networks", Proc. of 5th Intl. Conf. on Telecommunication Systems - Modeling and Analysis, March 20-23, 1997

[11] W. Mangione-Smith: "LPMAC - A Low Power Medium Access Control Protocol For Portable Multi-media Systems", Proc. of 3rd Intl. Workshop on Mobile Multimedia Comm., MoMuC-3, Sept. 25-27, 1996

[12] M. Zorzi, R.R. Rao: "Error Control and Energy Consumption in Wireless Communication Channels", IEEE Transaction on Computers (special issue on Mobile Computing), Mar. 1997

[13] P. Lettieri, Christina Fragouli, M. B. Srivastava: "Low Power Error Control for Wireless Links", Proc. of MobiCom'97, Sept. 1997, Budapest, Hungary, pp. 139 - 150

[14] G. D. Forney, Jr.: "Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference", IEEE Transactions on Information Theory, 1972, pp. 363-378

[15] M. Zorzi, R.R. Rao: "Energy Management in Wireless Communication", Proc. of 6th WINLAB workshop, pp.47, New Brunswick, NJ, USA, Mar. 1997

[16] A. Festag, J. Weinmiller, A. Wolisz: "Interconnection of Wireless Cells - a Multicast-based Approach" accepted at ICUPC '97 IEEE International Conference on Universal Personal Communications, San Diego, California, U.S.A. October 12 - 17, 1997

[17] David F. Bantz, Frederic J. Bauchot: "Wireless LAN Design Alternatives", IEEE Network Magazine, March/April, 1994

[18] Z. J. Haas: "Mobile-TCP: An Asymmetric Transport Protocol Design for Mobile Systems", Proc. of 3rd INTERNATIONAL WORKSHOP ON MOBILE MULTIMEDIA COMMUNICATIONS, Princeton, New Jersey, September 25-27, 1996

[19] E. Ayanoglu et al.: "AIRMAIL: A link-layer protocol for wireless networks", Wireless Networks, J.C. Baltzer AG, Science Publishers, vol 1, pp. 47-60

[20] “Inter Access Point Protocol, Revision 1.2,”, Aironet Wireless Communications Inc., Lucent Technologies, Digital Ocean, August 1996

[21] Jost Weinmiller, Hagen Woesner, Jean-Pierre Ebert, Adam Wolisz: ”Analyzing the RTS/CTS Mechanism in the DFWMAC Media Acces Protocol for Wireless LAN’s”, Proc. of IFIP TC6 Workshop Personal Wireless Communications (Wireless Local Access), April 1995, pp. 117-130

[22] J.Weinmiller, H. Woesner, A. Wolisz: ”Analyzing and Tuning the Distributed Coordination Function in the IEEE 802.11 Draft Standard”; June 1995 MASCOT ‘96; Feb. 1996, San Jose, California

[23] F. Cali, M.Conti, E. Gregori: “IEEE 802.11 Wireless LAN: Capacity Analysis and Protocol Enhancement”, Proc. of INFOCOM 1998, San Francisco, April 1998

[24] PTOLEMY; www-site: http://www.ptolemy.eecs.berkeley.edu, Copyright 1990-96 The Regents of the University of California

[25] R. Caceres, V. Padmababhan: ”Fast and Scalable Handoffs for Wireless Internetworks”; Proc. of MobiCom, November 1996

[26] J. Weinmiller, M. Schläger, A. Festag, A. Wolisz: ”Performance Study of Access Control in Wireless LANs - IEEE 802.11 DFWMAC and ETSI RES 10 HIPERLAN” in Mobile Networks and Applications Baltzer Science Publishers, July 97

[27] M. Tanabe, S. Hakomori, U. Inoue: ”Information-providing Mechanism Combining Broadcast and On-demand Modes in Mobile Computing Environments”, Proc. of 6th WINLAB workshop, pp.47, New Brunswick, NJ, USA, Mar. 1997

[28] A. DeSimone, M. C. Chuah, O.C. Yue: ”Throughput Performance of Transport-Layer Protocols over Wireless LANs”, Proceedings of IEEE GLOBECOM ‘93, Houston, 1993