IEEE 802.11N: ENHANCEMENTS FOR HIGHER THROUGHPUT IN WIRELESS LANS

YANG XIAO, THE UNIVERSITY OF MEMPHIS

(d) MAC level version

(e) MAC level version

(f) MAC level version

a single PHY frame
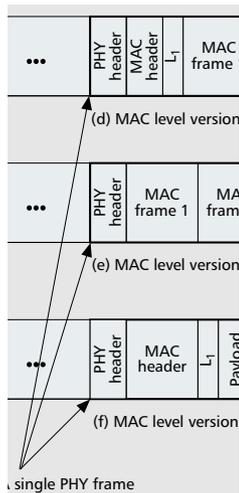
The authors introduce a new standardization effort, IEEE 802.11n, an amendment to IEEE 802.11 standards that it is capable of much higher throughputs, with a maximum throughput of at least 100Mb/s, as measured at the medium access control (MAC) data service access point.

ABSTRACT

This article introduces a new standardization effort, IEEE 802.11n, an amendment to IEEE 802.11 standards that it is capable of much higher throughputs, with a maximum throughput of at least 100 Mb/s, as measured at the medium access control data service access point. The IEEE 802.11n will provide both physical layer and MAC enhancements. In this article we introduce some PHY proposals and study the fundamental issue of MAC inefficiency. We propose several MAC enhancements via various frame aggregation mechanisms that overcome the theoretical throughput limit and reach higher throughput. We classify frame aggregation mechanisms into many different and orthogonal aspects, such as distributed vs. centrally controlled, ad hoc vs. infrastructure, uplink vs. downlink, single-destination vs. multi-destination, PHY-level vs. MAC-level, single-rate vs. multirate, immediate ACK vs. delayed ACK, and no spacing vs. SIFS spacing.

INTRODUCTION

Wireless local area networks (WLANs) are becoming more popular and increasingly relied on. The IEEE 802.11 WLAN is accepted as a complementary technology to high-speed IEEE 802.3 (Ethernet) for portable and mobile devices. One reason for such success is that it keeps increasing data transmission rates while maintaining a relatively low price. The IEEE 802.11, 802.11b, and 802.11a/g specifications provide up to 2 Mb/s, 11 Mb/s, and 54 Mb/s data rates [1, 2], respectively. Furthermore, the IEEE 802.11 Working Group is pursuing IEEE 802.11n, an amendment for higher throughput and higher speed enhancements. Different from the goal of IEEE 802.11b/.11a/.11g, i.e., to provide higher-speed data rates with different physical layer (PHY) specifications, IEEE 802.11n aims at higher throughput instead of higher data rates with PHY and medium access control (MAC) enhancements.

The IEEE 802.11 MAC employs a mandatory contention-based channel access function called the distributed coordination function (DCF) and an optional centrally controlled channel access function, the point coordination function (PCF) [1]. The DCF adopts carrier sense multiple access with collision avoidance (CSMA/CA) with binary exponential backoff, and the PCF adopts a polling mechanism. To support MAC-level quality of service (QoS), the IEEE 802.11 Working Group is currently working on the standardization of IEEE 802.11e. The IEEE 802.11e MAC employs a channel access function called the hybrid coordination function (HCF), which includes contention-based channel access, enhanced distributed channel access (EDCA), and contention-free centrally controlled channel access, HCF controlled channel access (HCCA).

To provide better QoS, especially for multimedia applications, increasing data rates is also highly desirable. The rationale is the same as Ethernet, which dramatically increases data rates from 10/100 Mb/s to 10 Gb/s. Data-rate-intensive applications exist such as multimedia conferencing, MPEG video streaming, consumer applications, network storage, file transfer, and simultaneous transmission of multiple HDTV signals, audio, and online gaming. Furthermore, there is a great demand for higher-capacity WLAN networks in the market in such areas ass hotspots, service providers, and wireless backhaul. Therefore, increasing data rates is crucial, and the IEEE 802.11 Working Group was seeking higher data rates over 100 Mb/s for IEEE 802.11a extension [2]. However, we proved that a theoretical throughput limit exists due to MAC and PHY overhead [2]. In other words, the theoretical throughput limit, about 75 Mb/s for IEEE 802.11a with a payload size of 1500 bytes [2], upper bounds any obtained throughput even when the data rate goes infinitely high. Therefore, increasing transmission rate cannot help a lot. Both reducing overhead and pursuing higher data rates are therefore necessary and important [2]. In July 2002 the IEEE 802.11 High Throughput Study Group (HTSG) was established, emphasizing higher throughput for data rates over 100 Mb/s in WLANs. The first official meeting of the IEEE 802.11n Task Group took place in September 2003, replacing the IEEE High Throughput Study Group (HTSG). The scope of IEEE 802.11n is to define an amendment to the IEEE 802.11 standards to enable much higher throughputs, with a maximum

throughput of at least 100 Mb/s, as measured at the MAC data service access point (SAP). IEEE 802.11n will provide both PHY and MAC enhancements. Note that even though IEEE 802.11e also provides some mechanisms for efficient MAC enhancements such as Direct Link Protocol and Block Acknowledgment Protocol, its major goal is still to provide QoS services, whereas the goal of IEEE 802.11n is to provide higher throughput via PHY and MAC enhancements.

In this article we first introduce the history and current status of IEEE 802.11n, as well as some PHY proposals. Then we study the theoretical throughput limit and provide an overhead analysis for IEEE 802.11, and compare this aspect with HIPERLAN/2. Finally, we propose several MAC enhancements via various frame aggregation mechanisms. We adopt the original IEEE 802.11 MAC in this article, but the mechanisms can easily be applied to IEEE 802.11e.

IEEE 802.11N

In this section we provide an up-to-date survey on the efforts to produce the IEEE 802.11n standard, including its history and current status. The IEEE 802.11n standard process has three phases: phase 1 is the preparation stage from January to September 2002; phase 2 was the of IEEE 802.11 HTSG from September 2002 to September 2003; phase 3 is the IEEE 802.11n Task Group (TGn), which began in September 2003 and is expected to finish in March 2007.

PHASE 1: PREPARATION

The first formal presentation at IEEE 802 meetings about 802.11a higher data rate extension was at the IEEE 802.11 interim meeting at Dallas, Texas, in January 2002 [2]. In this presentation Jones *et al.* described the high demand for data rates over 100 Mb/s through IEEE 802.11, and described some potential approaches to achieve higher data rates: modulation and coding enhancements, spatial diversity techniques, spatial multiplexing, and double bandwidth solutions with underlying IEEE 802.11a waveforms. Later on at the meeting, a straw poll for a call for interest on 802.11a higher rate extension was conducted in the IEEE 802.11 working group, and received tremendous interest among hundreds of committee members.

At the St. Louis IEEE 802 plenary meeting in March 2002, we provided a throughput analysis for higher data rates over 100 Mb/s [2]. Tzannes *et al.* proposed a bit-loading (BL) approach [2]. One of the drawbacks is that BL may require feedback from the receiver to the transmitter, and the communication from the receiver to the transmitter must happen faster than channel changes. Hori *et al.* compared four different potential approaches: the double clock rate approach, the double subcarrier number approach, the 4096-quadrature amplitude modulation (QAM)-orthogonal frequency-division multiplexing (OFDM) approach, and the OFDM/space-division multiplexing (SDM) (multicarrier multiple-input multiple-output [MIMO]) system approach [2]. Coffey suggested some criteria for higher data rates, and that

higher data rates should emphasize throughput with consideration for backward compatibility rather than data rate [2].

At the Sydney, Australia, IEEE 802 interim meeting in May 2002, we showed that a theoretical throughput upper limit exists for IEEE 802.11 protocols [2]. Therefore, increasing transmission rate cannot help much. Both reducing overhead and pursuing higher data rates are therefore necessary and important [3]. The IEEE 802.11 HTSG was established in July 2002 emphasizing higher throughput for higher data rates over 100 Mb/s WLANs.

PHASE 2: IEEE 802.11 HTSG

The IEEE 802.11 HTSG began operations in September 2002 and ended in September 2003. During this phase, a Project Authorization Request (PAR) and Five Criteria for Standards Development were established.

The scope of the MAC and PHY enhancements assume a baseline specification to support higher throughput. The amendment seeks to improve the peak throughput to at least 100 Mb/s, measured at the MAC data SAP. This represents an improvement of at least four times the throughput obtainable using existing 802.11 systems. The highest throughput mode shall achieve a spectral efficiency of at least 3 b/s/Hz. The Task Group (IEEE 802.11n) will undertake the following steps:
- Identify and define usage models, channel models, and related MAC and application assumptions.
- Identify and define evaluation metrics that characterize the important aspects of a particular usage model.

Initial usage models include hotspot, enterprise, and residential. Evaluation metrics include throughput, range, aggregate network capacity, power consumption (peak and average), spectral flexibility, cost/complexity flexibility, backward compatibility, and coexistence.

The Five Criteria for Standards Development are:
- **Broad market potential**: It shall have a broad market potential; that is, broad sets of applicability, multiple vendors and numerous users, and balanced costs (LAN vs. attached stations).
- **Compatibility**: Keeping the MAC SAP interface the same as for the existing 802.11 standards is required for compatibility. New enhancements shall be defined in a format and structure consistent with existing 802.11 standards.
- **Distinct identity**: Each IEEE 802 standard shall have a distinct identity from other IEEE 802 standards.
- **Technical feasibility**: Those introduced in phase 1 and later parts of this subsection can provide technical feasibility. Furthermore, there are currently reliable WLAN solutions.
- **Economic feasibility**: Economic feasibility includes known cost factors, reasonable cost for performance, and consideration of installation costs.

Next, we introduce some additional proposals for technical feasibility. In [4] the authors proposed exploring space diversity through multiple

antennae other than frequency diversity via bit interleaved coded modulation, and claimed that OFDM is very well suited for use with multiple antennae, for example, as an optional mode in IEEE 802.16, with the cost of an additional antenna and a radio frequency (RF) front-end. In [5] the authors proposed a combined scheme of BL and trellis-coded modulation (TCM). BL is the process of modulating a different number of bits on each carrier based on the signal-to-noise ratio (SNR) of the carriers. IEEE 802.11a adopts an equal number of bits per carrier, which is the simplest form of BL. BL is better suited to a multipath channel. However, it requires feedback from the receiver to the transmitter, and the communication from the receiver to the transmitter must happen faster than channel changes. On the other hand, TCM combines coding and modulation functions to provide improved performance. Coded modulation schemes combine with BL to encode all information bits. The advantages include:

- It does not have a preset maximum data rate.
- It is optimally suited to a multipath channel.
- It is based on mature and widely understood technology, such as asymmetric digital subscriber line (ADSL).
- It requires a relatively small standardization effort.

In [6] the author proposed a MIMO-OFDM solution. In [7] the authors claimed that 250 Mb/s data rate is achievable. In [8] the authors showed some experimental results based on MIMO-OFDM for high-throughput WLANs. In [9] the author proposed using smart antennas to improve SNR, increase coverage/range and data rate, reduce interference and multipath, and increase network capacity and battery life.

PHASE 3: IEEE 802.11N TGN

The first official meeting of the IEEE 802.11n Task Group took place September 2003 in Singapore. The IEEE 802.11n standard is planned to be published in March 2007. The TGn will further go through the following steps: establishing the proposal selection process and criteria, call for proposals, combination of proposals, several letter ballots, standard approval, and finally standard publication. We will see more contributions in future IEEE 802.11n meetings. So far, most contributions in IEEE 802.11n meetings focus on PHY enhancements. Instead, this article serves a good purpose in discussing MAC enhancements.

THEORETICAL LIMIT AND OVERHEAD ANALYSIS OF THE IEEE 802.11 MAC

The achievable maximum throughput (MT) can be met when the system is in the best case scenario:

- The channel is ideal, without errors.
- At any transmission cycle, there is one and only one active station that always has a frame to send, and other stations can only accept frames and provide acknowledgments (ACKs).

The throughput upper limit (TUL) [2] is defined as the maximum throughput when the raw data rate goes infinitely high.

As indicated in [3], overhead is the major fundamental issue for inefficient MAC, and it includes headers (MAC header, frame check sequence [FCS], and PHY header), interframe spaces (IFSs), backoff time, and ACKs. Define overhead as the difference between data rate and throughput, and define normalized overhead as overhead divided by data rate. We further assume that all higher data rates are compatible with IEEE 802.11a. Let T_{slot}, T_{SIFS}, T_{DIFS}, and CW_{min} denote a slot time, a short IFS (SIFS) time, a differentiated IFS (DIFS) time, and the minimum backoff contention window size, respectively. Let T_P and T_{PHY} denote transmission times of a physical preamble and a PHY header, respectively. Let T_{DATA} and T_{ACK} denote transmission times of a data frame and an ACK, respectively. Let L_{DATA} denote the payload length in bytes.
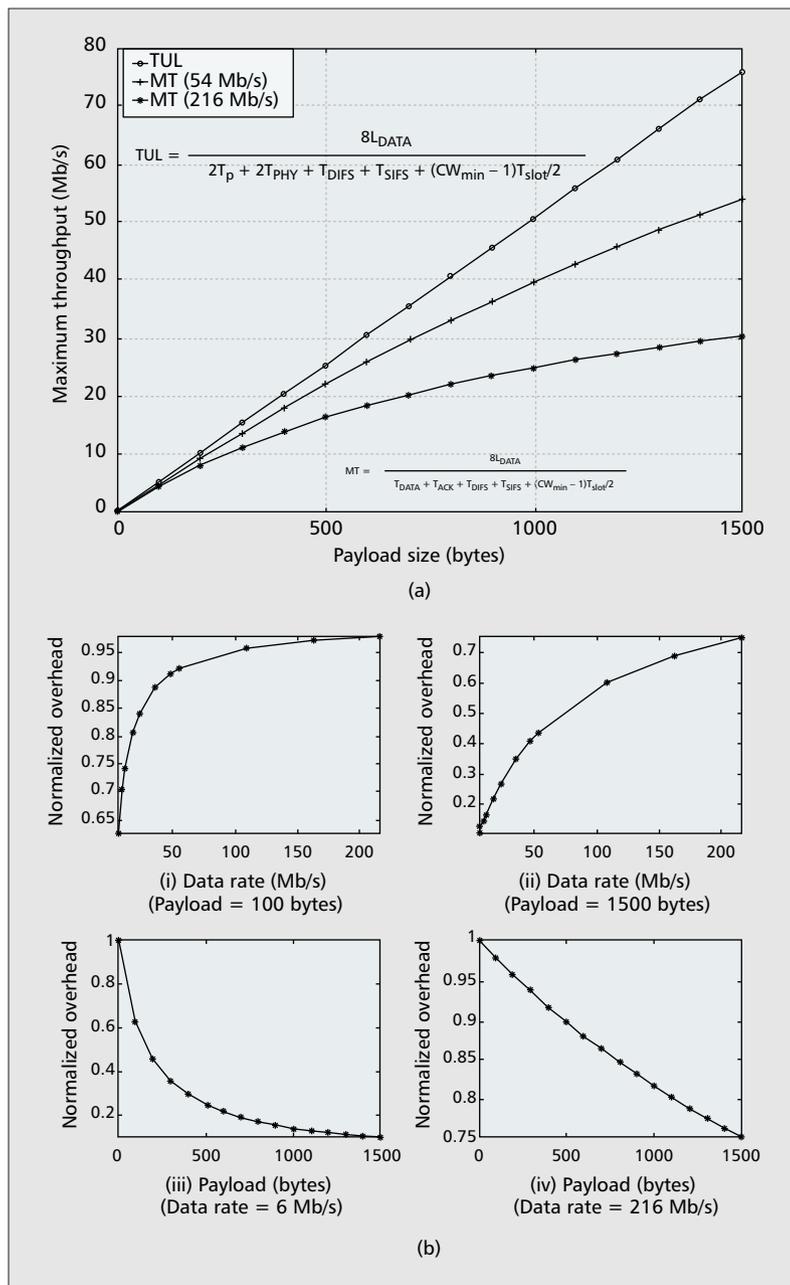
■ **Figure 1**. *a) The MT and TUL for IEEE 802.11a; b) normalized overhead vs. data rate and payload.*
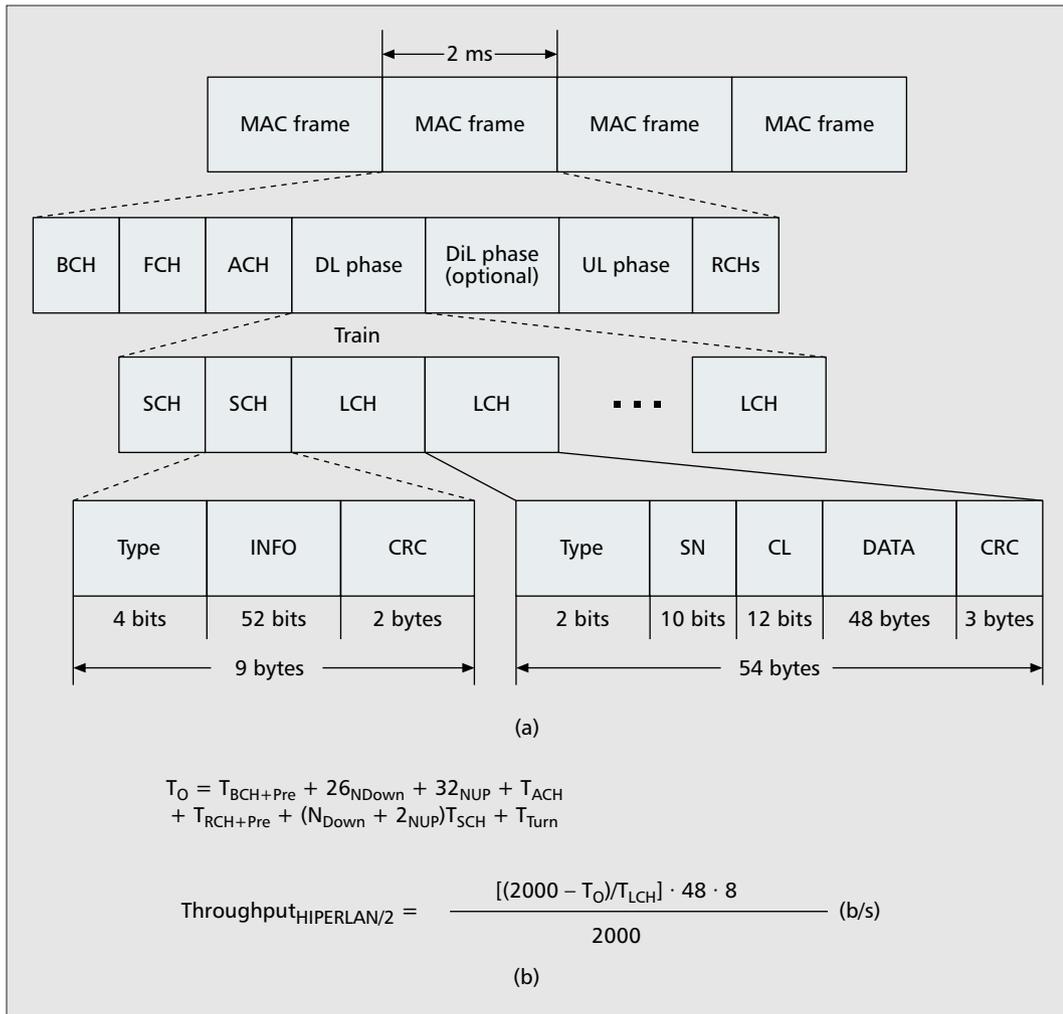
Figure 2. *a) HIPERLAN/2 MAC frame; b) equations for overhead and throughput.*

According to our previous contributions in [2], the MT and TUL are shown in Fig. 1a.

Figure 1a shows the MT and TUL for IEEE 802.11a. As illustrated in the figure, the TUL upper bounds the MT at a 54 Mb/s data rate and the MT at a 216 Mb/s data rate. When the payload size is 1500 bytes, the TUL is about 75.24 Mb/s. The existence of the TUL shows that by simply increasing the data rate without reducing overhead, the enhanced throughput is bounded even when the data rate goes infinitely high. In other words, reducing overhead is necessary for IEEE 802.11 standards to achieve higher throughput.

Figure 1b(i) and (ii) show normalized overhead vs. data rate. The normalized overhead increases as the data rate increases. The normalized throughput almost reaches 1 after 180 Mb/s when the payload size is 100 bytes. The normalized throughput reaches 70 percent after 180 Mb/s when the payload size is 1500 bytes. Figure 1b(iii) and (iv) show normalized overhead vs. payload size. The normalized overhead decreases as the payload size increases. The normalized throughput almost reaches 1 when the payload size is very small.

In summary, the normalized overhead is extremely large when either the data rate is high or the frame is small.

HIPERLAN/2

The wireless LAN standards, European Telecommunications Standards Institute (ETSI) Broadband Radio Access Network (BRAN) HIPERLAN/2 [10] and IEEE 802.11a/11g, offer transmission rates up to 54 Mb/s in the 5 GHz or 2.4 GHz band. In this section we compare IEEE 802.11a and HIPERLAN/2 in terms of throughput upper limit. The two standards differ primarily in the MAC layer. The actual throughputs achieved are also highly dependent on MAC protocols. HIPERLAN/2 employs centralized control, where a scheduler at an AP allocates resources in a MAC frame. Another difference in MAC protocols is the packet length adopted: HIPERLAN/2 adopted fixed length packets, and IEEE 802.11 adopted variable length packets.

In HIPERLAN/2 a MAC frame is transmitted in a period of 2 ms (Fig. 2a). Each MAC frame comprises time slots for broadcast control (BCH), frame control (FCH), access feedback control (ACH), data transmission in downlink (DL), direct link (DiL), and uplink (UL) phases, and random channels (RCHs). DL and UL are used when data has to be transmitted. DL, UL, and DiL phases consistent of two types of protocol data units (PDUs): the short transport chan-

To overcome overhead of the IEEE 802.11 MAC, we propose several efficient MAC enhancements, in which we adopt frame aggregation concept. The idea of frame aggregation is to aggregate multiple MAC/PHY frames/payloads into a single (or approximately single) transmission.

nel (SCH) and long transport channel (LCH). SCHs are for control data and have a size of 9 bytes. LCHs are for normal data and have a size of 54 bytes, including 6 bytes overhead and 48 bytes payload. A train of SCH and LCH packets is transmitted in DL and/or UL. The duration of BCH is fixed; others are not. When transmitted, there is a physical preamble for the whole MAC frame, DL, DiL, UL, and RCH, respectively. We do not consider DiL since it is optional and not important for conclusions.

Let $T_{BCH+Pre}$, T_{FCH}, T_{ACH}, $T_{RCH+Pre}$, T_{SCH}, T_{LCH}, T_{DPre}, and T_{UPre} denote the transmission times for a BCH with the preamble, FCH, ACH, RCH with preamble, SCH, LCH, preamble in DLs, and preamble in ULs, respectively. Let T_{CD} and T_{CU} denote the transmission times for control signal in DL (an SCH for ACK per session) and control signal in UL (the SCH for ACK and resource reservation per session), respectively. Let T_{GU}, T_{Turn}, and T_O denote guard time in UL, radio turnaround time, and total overhead time, respectively. Data rates are 6, 9, 12, 18, 27, 36, and 54 Mb/s, and the corresponding T_{LCH} are 72, 48, 36, 24, 16, 12, and 8 µs, respectively. Let N_{Down} and N_{Up} denote the number of sessions in UL and DL, respectively. From [10] we have $T_{FCH} = [(N_{Down} + N_{Up})/2]\,36$ µs, $T_{CD} = N_{Down}T_{SCH}$, $T_{CU} = 2N_{Up}T_{SCH}$, $T_{DPre} = N_{Down}\,8$ µs, $T_{UPre} = N_{Up}\,12$ µs, and $T_{GU} = N_{Up}\,2$ µs. We have overhead and throughput shown in Fig. 2b. From [10] we have $T_{BCH+Pre} = 36$ µs, $T_{ACH} = 12$ µs, $T_{RCH+Pre} = 28$ µs, and $T_{Turn} = 12$ µs. We obtain overhead and throughput shown in Fig. 2b. When both the data rate and control rate go infinitely high, we have $T_{SCH} = 0$ and $T_{LCH} = 0$. Furthermore, when the data rate goes infinite, the total overhead T_O will be a fixed number, and *Throughput_HIPERLAN/2* will be infinitely large (∞). In other words, the throughput upper limit of HIPERLAN/2 is ∞.

IEEE 802.11 employs a protocol similar to the stop-and-wait protocol so that every packet is acknowledged. In each transmission cycle there is a fixed overhead time, including spacing, ACK, and other overhead, independent of the data rate. For a given payload size, no matter how high the data rate is, the transmission time is larger than the fixed overhead time. Therefore, IEEE 802.11a has a throughput limit, which is 75.24 Mb/s when the payload size is 1500 bytes. On the other hand, HIPERLAN/2 employs a MAC frame (transmitted in 2 ms) in which a train of LCHs can be transmitted within 2 ms. As the data rate goes infinitely high, the number of LCHs that can be potentially transmitted is infinite. Therefore, HIPERLAN/2's throughput upper limit is ∞. In other words, HIPERLAN/2 does not have a throughput upper limit. The conclusion is that HIPERLAN/2 is much more scalable than IEEE 802.11 for much higher data rates.

IEEE 802.11 MAC ENHANCEMENTS

To overcome the overhead of the IEEE 802.11 MAC, we propose several efficient MAC enhancements in which we adopt the frame aggregation concept. The idea of frame aggregation is to aggregate multiple MAC/PHY frames/payloads into a single (or approximately single)
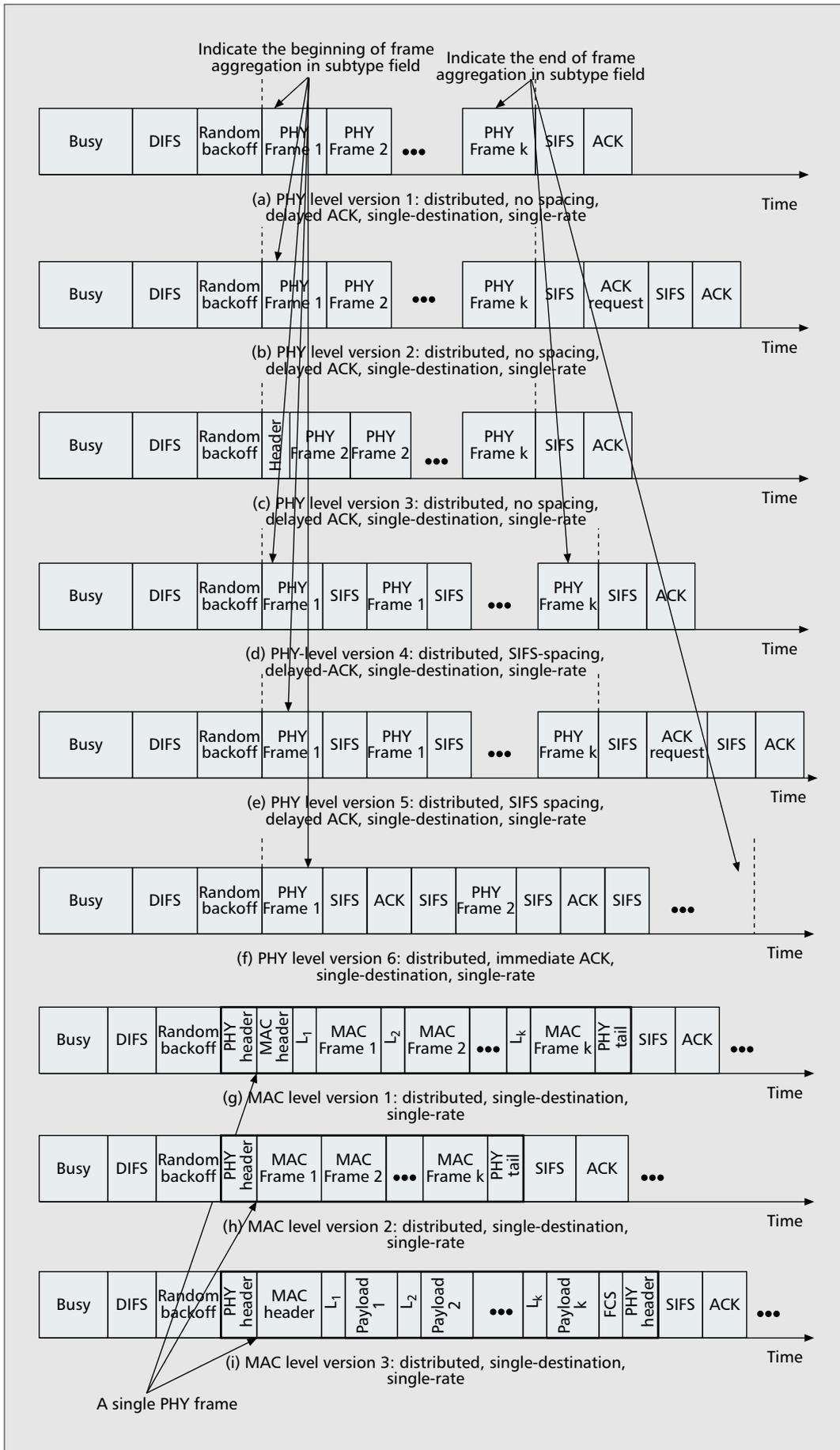
transmission. We classify frame aggregation mechanisms into many different and orthogonal aspects:

- **Distributed vs. centrally controlled**: Frame aggregation can be used under both the contention-based DCF and the contention-free PCF. The former is a distributed mechanism, and the latter is centrally controlled.
- **Ad hoc vs. infrastructure**: There are two types of 802.11 networks: infrastructure (BSS) in which an AP is present, and ad hoc (IBSS) in which an AP is not present. Frame aggregation can be used in both ad hoc and infrastructure networks.
- **Uplink vs. downlink**: In infrastructure networks, frame aggregation can be used by both UL and DL transmissions. UL transmissions are those from stations to the AP, and DL transmissions are those from the AP to stations.
- **Single-destination vs. multi-destination**: Under some frame aggregation mechanisms, frames can be aggregated only if they have the same destination (MAC/PHY) addresses. Under other proposed frame aggregation mechanisms, frames can be aggregated even if they have different destination (MAC/PHY) addresses.
- **PHY-level vs. MAC-level**: Frames can be aggregated at both the PHY and MAC levels. At the PHY level, contents of PHY frames retain integrity. At the MAC level, PHY frames are changed, and MAC frames may or may not be changed. However, payloads of MAC frames cannot be changed.
- **Single-rate vs. multirate**: Aggregated frames can use the same transmission rate, and can use different transmission rates for multi-destination frame aggregation.
- **Immediate ACK vs. delayed ACK**: Aggregated frames can be acknowledged via separate ACK frames immediately following a SIFS time, or a single delayed group ACK frame can be adopted after aggregated frame transmissions.
- **No spacing vs. SIFS spacing**: Under delayed ACK frame aggregation, two consecutive aggregated PHY frames may be separated by nothing or a SIFS time.

Frame aggregation mechanisms have many benefits. First of all, since transmitting longer frames may lead to better throughput than transmitting shorter frames, by adopting these mechanisms the system can achieve the throughput of transmitting longer frames. The second and most important benefit is that these mechanisms can reduce overhead. Without these mechanisms, each frame transmission needs a separate set of overhead (headers, IFSs, backoff time, and/or ACKs). With these mechanisms, instead of several sets of overhead for different frames, only one set of overhead will be used. Finally, these mechanisms can reduce the average delay. Without these mechanisms, the second or a later frame is transmitted at a much later time. With these mechanisms, it is transmitted at almost the same or earlier.

One issue is how long the total length of aggregated frames should be. One possible solution is that the number of aggregated frames should not be larger than a threshold, and the

Figure 3. *Distributed, single-destination, and single-rate frame aggregation.*

The figure includes the following labeled diagrams:

Indicate the beginning of frame aggregation in subtype field

Indicate the end of frame aggregation in subtype field

(a) PHY level version 1: distributed, no spacing, delayed ACK, single-destination, single-rate

(b) PHY level version 2: distributed, no spacing, delayed ACK, single-destination, single-rate

(c) PHY level version 3: distributed, no spacing, delayed ACK, single-destination, single-rate

(d) PHY-level version 4: distributed, SIFS-spacing, delayed-ACK, single-destination, single-rate

(e) PHY level version 5: distributed, SIFS spacing, delayed ACK, single-destination, single-rate

(f) PHY level version 6: distributed, immediate ACK, single-destination, single-rate

(g) MAC level version 1: distributed, single-destination, single-rate

(h) MAC level version 2: distributed, single-destination, single-rate

(i) MAC level version 3: distributed, single-destination, single-rate

A single PHY frame

Frame aggregation mechanisms have many benefits. Without these mechanisms, each frame transmission needs a separate set of overhead. With these mechanisms, instead of several sets of overhead for different frames, only one set of overhead will be used.

In single-destination approaches, frames can be aggregated if they are available, and have the same source and destination addresses. The total length of aggregated frames should be smaller than a threshold, which is called aggregation threshold.

total length of aggregated frames should be smaller than another threshold, which is smaller than or equal to the fragmentation threshold. The purpose of these mechanisms is not to build a huge frame, but a reasonably sized frame since huge frames may have a bad effect on fairness and/or efficiency. Furthermore, frame aggregation is not a reversed mechanism of fragmentation. In fact, the proposed aggregated mechanisms require that the total length of the aggregated frames be smaller than the fragmentation threshold. Therefore, there will be no aggregated frame that was originally generated by a previous fragmentation mechanism. On the other hand, an aggregated frame will not be fragmented since the total length is smaller than the fragmentation threshold. Next, we discuss some frame aggregation mechanisms.

DISTRIBUTED, SINGLE-DESTINATION, AND SINGLE-RATE FRAME AGGREGATION

Distributed, single-destination, and single-rate frame aggregation mechanisms are shown in Fig. 3, where Figs. 3a, 3b, 3c, 3d, 3e, and 3f are PHY-level frame aggregation methods, and Figs. 3g, 3h, and 3i are MAC-level frame aggregation methods. At the PHY level contents of PHY frames remain unchanged, and at the MAC level PHY frames are changed; MAC frames may or may not be changed. Note that the payloads of MAC frames cannot be changed. In single-destination approaches, frames can be aggregated if they are available, and have the same source and destination addresses. The total length of aggregated frames should be smaller than a threshold, called the *aggregation threshold*.

Figure 3a shows that k PHY frames are sent one by one with no spacing between PHY frames. The first aggregated frame (PHY frame 1) should indicate the beginning of the frame aggregation in one of the subtype fields, and the last aggregated frame (PHY frame k) should indicate the end of frame aggregation, followed by a SIFS time and an ACK from the destination station. The destination can identify the PHY boundaries via PHY preambles. The ACK frame should have a bitmap field, in which each bit can be used to acknowledge one frame: a bit value of 1 stands for successfully receiving a frame; it is 0 otherwise.

Figure 3b has minor differences from Fig. 3a as follows. Instead of indicating the end of frame aggregation in the last aggregated frame (PHY frame k), an ACK request frame should send directly from the source, and the destination respond with an ACK frame after SIFS time. This scheme adds additional overhead, but the system becomes more robust since if the last aggregated frame (PHY frame k) in the scheme in Fig. 3a is lost, the destination does not obtain an explicit indication that an ACK needs to be sent back.

Figure 3c shows another variant of Fig. 3a: instead of using the first aggregated frame to indicate the beginning of the frame aggregation, another header frame is transmitted earlier in which more information is included such as the number (1 byte) of aggregated frames to follow and the total length/duration (2 bytes). The drawback is that an additional header PHY frame is needed.

Figure 3d is another variant of Fig. 3a, and Fig. 3e is a variant of Fig. 3b. In both Figs. 3d

and 3e, a SIFS time is used to separate PHY frames. A drawback is that additional overhead of SIFS time is needed. An advantage is that the schemes make PHY frames' boundaries clearer.

Figure 3a, 3b, 3c, 3d, and 3e all adopt delayed ACK, whereas Fig. 3f adopts immediate ACK. The drawback of Fig. 3f is that it has more overhead. One advantage is that it becomes more robust under a noisy channel.

Roughly speaking, comparing schemes in Figs. 3a–3f, we observe that if a scheme has more overhead, it becomes less efficient, but is more robust in terms of error control, explained as follows. The scheme in Fig. 3f is the most inefficient of the six schemes in Figs. 3a–3f, but the most robust for error control: if an ACK has not been received within SIFS+ACK time, frame aggregation will be stopped, and all successfully transmitted frames no longer need to be retransmitted. The schemes in Fig. 3d and 3e are more efficient than that in Fig. 3f due to delayed ACK, but error control is less robust than with the scheme in Fig. 3f: during aggregation frame transmissions, there is no way for the source to know that one or more frames get corrupted. The schemes in Figs. 3d and 3e are less efficient than the schemes in Figs. 3a–3c, but more robust in terms of distinguishing frame boundaries. One drawback of the schemes in Fig. 3a, 3c, and 3d is that if the ACK is corrupted, all the aggregated frames are retransmitted, whereas in the schemes in Figs. 3b and 3e, if either the ACK request frame or the ACK frame is corrupted, the source station can send the ACK request frame again after a timeout period.

Figures 3g, 3h, and 3i show three MAC-level frame aggregation schemes, and the order of these schemes is the order of increasing efficiency, but also the order of increasing complexity. Figure 3g shows that several MAC frames are aggregated together with an additional MAC header and several length fields ($L_1, …, L_k$), and embedded in a PHY frame. The scheme in Fig. 3h becomes more efficient than that in Fig. 3g since there are no additional MAC header and length fields, but the first aggregated frame (MAC frame 1) needs to be changed to indicate that this is an aggregated frame, and the duration field also needs to change. In contrast, MAC frames of the scheme in Fig. 3g remain intact. For the scheme in Fig. 3h, the destination station can easily identify boundaries of aggregated frames using the FCS and MAC duration field to calculate the length of the MAC frame. The scheme in Fig. 3i becomes more efficient than that in Fig. 3g since MAC headers and FCS are all removed, and a new header and FCS are added. In other words, the scheme in Fig.3i performs payload aggregation instead of frame aggregation. Therefore, we claim that the ordering of Figs. 3g, 3h, and 3i is that of increasing efficiency, but also of increasing complexity or processing time.

More variants of PHY- and MAC-level frame aggregation mechanisms can easily be designed.

DOWNLINK FRAME AGGREGATION

For infrastructure networks, in both the distributed (contention-based) DCF and the centrally controlled PCF, DL frame aggregation can be adopted. Note that all frame aggregation schemes in Fig. 3 can be used in UL transmis-
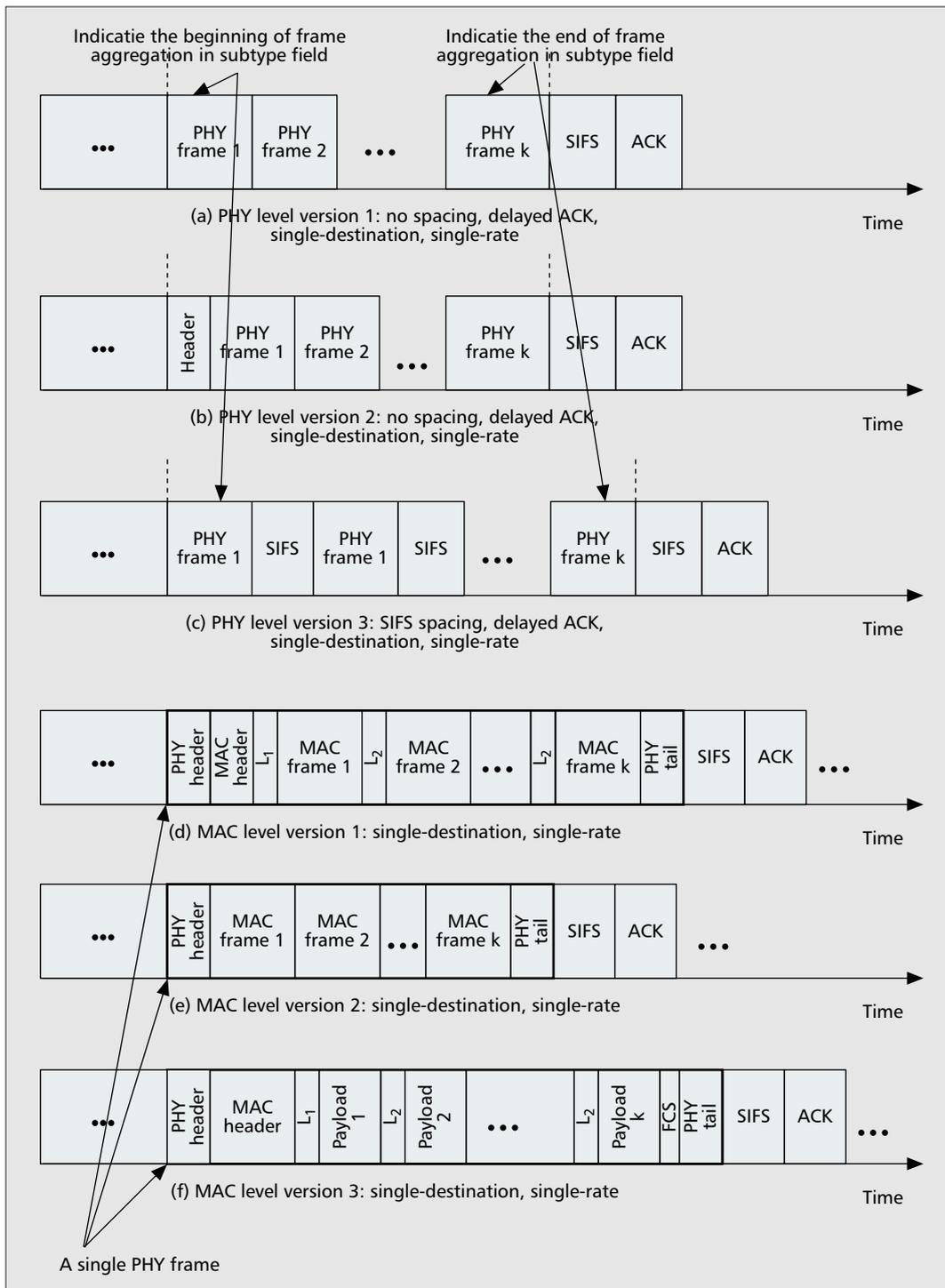
■ Figure 4. *Downlink frame aggregations.*

sions. Figure 4 shows some examples of DL frame aggregation schemes, which can be used in both the DCF and PCF from the AP to stations. Figure 4 is similar to Fig. 3 except that the frame aggregation is done in the AP in either the DCF or the PCF. Figures 4a, 4b, and 4c are PHY-level frame aggregation, whereas Figs. 4d, 4e, and 4f are MAC-level frame aggregation.
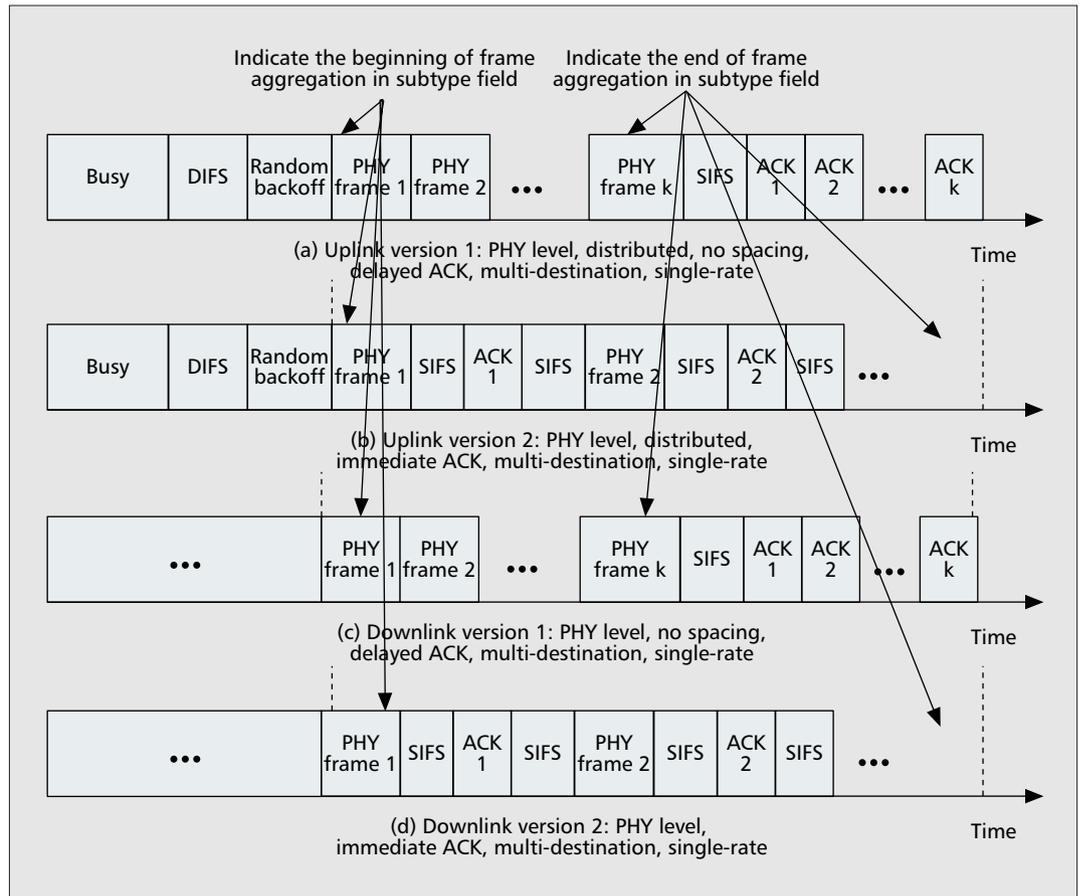
MULTI-DESTINATION FRAME AGGREGATION

Frame aggregation schemes in previous sections are all single-destination schemes. Figure 5 shows some multi-destination frame aggregation

schemes. Figures 5a and 5b are UL frame aggregation schemes in infrastructure networks or distributed frame aggregation schemes in ad hoc networks. Figures 5c and 5d are DL frame aggregation schemes in either the DCF or the PCF. In Figs. 5a and 5c, the aggregated frames can be for different destinations, and each destination station sendS a separate ACK per received frame with the same order of receiving frames. In Figs. 5b and 5d, the aggregated frames can be for different destinations, and each destination station sends an ACK immediately after receiving a frame after a SIFS time.

■ **Figure 5**. *Multi-destination frame aggregation.*

MULTIRATE OR CENTRALLY CONTROLLED FRAME AGGREGATION

Multi-destination frame aggregation schemes introduced in the previous section can easily be applied to multirate frame aggregation schemes, in which each aggregated PHY frame can adopt a different transmission rate for a different station based on current channel conditions between two stations. Multirate can be used for both DL and UL.

We introduced DL centrally controlled (PCF) frame aggregation earlier. For UL centrally controlled frame aggregation can easily be designed based on previous examples.

THROUGHPUT AND THROUGHPUT UPPER LIMIT

We study the maximum throughput (MT) and throughput upper limit (TUL) for the frame aggregation schemes. Without losing generality, we just use one frame aggregation scheme (i.e., the scheme in Fig. 3a) as an example in this section. For other frame aggregation schemes, similar results can easily be obtained.

Assume that all data frames are the same size, and at all times frames are concatenated. The MT and TUL for the scheme in Fig. 3a are given as

$$MT_a = \frac{8k * L_{DATA}}{kT_{DATA} + T_{ACK} + T_{DIFS} + T_{SIFS} + (CW_{\min} - 1)T_{slot}/2} \quad (1)$$

$$TUL_a = \frac{8k * L_{DATA}}{(k+1)(T_p + T_{PHY}) + T_{DIFS} + T_{SIFS} + (CW_{\min} - 1)T_{slot}/2} \quad (2)$$

Figure 6 shows the MT and TUL for the frame aggregation scheme and the original MAC. The data rate is 54 Mb/s, and we have $k = 2$. Figure 6 shows that the frame aggregation scheme has better MT and TUL.

CONCLUSIONS

In this article we introduce the history and current status of IEEE 802.11n, and many PHY enhancements for higher data rates. We identify overhead as the fundamental problem of MAC inefficiency. Increasing transmission rate alone cannot help a lot. The overhead is very large when either the data rate is high or the frame size is small. Therefore, new efficient MAC strategies are especially needed. We propose several MAC enhancements to reduce overhead via frame aggregation from various aspects such as distributed vs. centrally controlled, ad hoc vs. infrastructure, uplink vs. downlink, single-destination vs. multi-destination, PHY-level vs. MAC-level, single-rate vs. multirate,

immediate ACK vs. delayed ACK, and no spacing vs. SIFS spacing. A new throughput upper limit for the frame aggregation scheme is formulated. Our results show that performance is improved with frame aggregation schemes. We also compare IEEE 802.11 DCF with HIPERLAN/2 in terms of throughput upper limit, and show that HIPER-LAN/2 is much more scalable than IEEE 802.11 for much higher data rates.

REFERENCES

[1] IEEE 802.11, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification," Aug. 1999.
[2] Y. Xiao and J. Rosdahl, "Throughput and Delay Limits of IEEE 802.11," *IEEE Commun. Lett.*, vol. 6, no. 8, Aug. 2002, pp. 355–57.
[3] Y. Xiao and J. Rosdahl, "Performance Analysis and Enhancement for the Current and Future IEEE 802.11 MAC Protocols," *ACM Sigmobile Mobile Comp. and Commun. Rev.*, vol. 7, no. 2, Apr. 2003, pp. 6–19.
[4] M. Ghosh, X. Ouyang, and G. Dolmans, "On the Use of Multiple Antennae for 802.11," IEEE 802.11-02/180r0.
[5] M. Tzannes *et al.*, "Extended Data Rate 802.11a," IEEE 802.11-02/232r0.
[6] H. Yu, "HDR 802.11a Solution Using MIMO-OFDM," IEEE 802.11-02/294r1.
[7] G. Fettweis, G. Nitsche, and S. AG, "1/4 Gb/s WLAN," IEEE 802.11-02/320r0.
[8] A. Gorokhov *et al.*, "MIMO-OFDM for High Throughput WLAN: Experimental Results," IEEE 802.11-02/708r0.
[9] J. Regnier, "Benefits of Smart Antennas in 802.11 Networks," IEEE 802.11-03/025r0.
[10] ETSI, "Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Data Link Control (DLC) Layer; Part 1: Basic Transport Functions," Dec.1999.

BIOGRAPHIES

YANG XIAO [SM'04] (yangxiao@ieee.org) worked at Micro Linear as a MAC architect involved in IEEE 802.11 standard enhancement work before he joined the Department of Computer Science at The University of Memphis in 2002. He was a voting member of the IEEE 802.11 Working Group from 2001 to 2004. He currently serves as Editor-in-Chief for *International Journal of Security and Networks* and *International Journal of Sensor Networks*. He serves as an Associate Editor or on editorial boards of *International Journal of Communications Systems*, *Wireless Communications and Mobile Computing* (both, Wiley), *EURASIP Journal on Wireless Communications and Networking*, and

International Journal of Wireless and Mobile Computing. He has served as lead or sole guest editor for five journals during 2004–2005. He serves as a TPC member for many conferences, such as ICDCS, ICC, GLOBECOM, WCNC, ICCCN, PIMRC, and WMASH. He serves as a referee for many funding agencies, as well as a panelist for the U.S. National Science Foundation. His research areas include wireless networks and network security.

■ **Figure 6**. *Comparisons of MTs and TULs.*

EDITORS NOTE/ *(from page 2)*

networks, and how they can be integrated with satellites. The article by Wang *et al.* discusses issues related to congestion control in the space Internet, comparing window- and rate-based solutions in typical satellite environments. The article by Roy-Chowdhury *et al.* addresses the issue of security, which of course is of primary importance in wireless networks, including their satellite segments. The article by Albertazzi *et al.* discusses the physical layer design and performance of DVB-S2, one of the most promising standards for satellite broadcasting.

In addition, this issue also features an article by Li *et al.* that surveys the current state of power-aware routing protocols in ad hoc networks, and an article by Yang Xiao that describes PHY and MAC enhancements to the IEEE 802.11 standard for higher data rates.

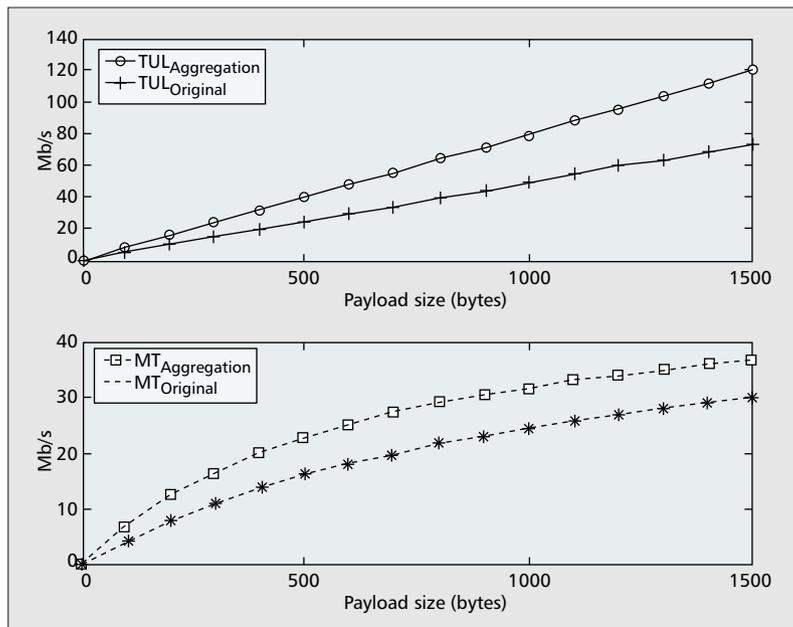I hope you will enjoy this issue, which is the last of my term, and wish you all a relaxing and refreshing holiday season and a fulfilling New Year.