I. Modeling and Analyzing MAC Frame Aggregation Techniques in 802.11n Using Bi-dimensional Markovian Model

Nazeeruddin Mohammad & Shahabuddin Muhammad
College of Computer Engineering and Science,
Prince Mohammad Bin Fahd University (PMU),
Al Khobar, Kingdom of Saudi Arabia
{nmohammad,smuhammad}@pmu.edu.sa

Abstract:
Increased expectations and demand for higher rates led to the development of new physical layer technologies in Wireless LANs. However, the current medium access control (MAC) needs to be improved to fully utilize higher physical-layer transmission rates. Several aggregation mechanisms have been recently proposed to improve the MAC layer performance of 802.11n. In this paper, we analyze some of the key aggregation mechanisms proposed. For analysis we adapted widely used Bianchi's analytical model and applied it for various aggregation techniques. We also compare the analytical details of various strategies and provide a unified analytical framework for continued research in this direction.

Keywords: 802.11n, Aggregation, CSMA/CA, MAC, WLANs

1 Introduction

Wireless LANs (WLAN) are becoming increasing ubiquitous because of the flexibility and the freedom they offer to the users. Their popularity has led to several improvements in the physical layer technologies for wireless networks. The latest IEEE 802.11n [1] standard has utilized MIMO (multiple-input and multiple-output) and OFDM (orthogonal frequency division multiplexing) techniques to achieve transmission rates upto 600 Mbps [2, 3]. However, several studies have shown that the improvements in the physical layer data rates will not be sufficient to improve the overall throughput [4]. The primary reason behind this is the overheads related to physical layer headers and contention time. These overheads do not decrease proportionately with the increase in physical data rates and dominate frame transmission time at higher physical data rates.

The throughput at MAC layer can be improved by mitigating frame overheads & contention time. One way of achieving this is by aggregating several frames in a single large frame, thereby minimizing channel idle time (fewer SIFS and backoffs) and frame overheads (fewer PHY headers). Few schemes [5, 6] tried to improve MAC efficiency by sending a train (burst) of frames after winning.

DCF contention window, and hence sharing contention overhead across multiple frames. In BlockAck strategy [7] a train of frames are transmitted without waiting for individual acknowledgments (ACK), and then the whole block is acknowledged (BACK) with single acknowledgment frame, thereby reducing the overhead due to ACKs and SIFS.

IEEE 802.11n Task Group (TGn) has adopted frame aggregation to improve the MAC layer efficiency and throughput. Frame aggregation schemes attempt to improve MAC efficiency by reducing waiting time during CSMA/CA back-off period for successive frame transmissions, and minimizing the transmission time for preamble and frame headers [8]. IEEE 802.11n TGn has defined two frame aggregation schemes, namely aggregate MAC Service Data Unit (A-MSDU) and aggregate MAC Protocol Data Unit (A-MPDU). A-MSDU mechanism joins several MAC Service Data Units (MSDUs) to form a single big MAC Protocol Data Unit (MPDU). On the other hand, A-MPDU concatenates several MSDUs (each with their own MAC header and FCS) to form Physical-layer Service Data Unit (PSDU). However, transmitting large frames is not encouraged in error-prone channels because single bit-error causes all frames to be retransmitted. Aggregation with Fragment Retransmission (AFR) [4] scheme tries to address this issue by providing mechanisms for partial retransmission of affected frames.

In addition to aggregation, 802.11n standard specifies bidirectional data transfer method over a single transmission opportunity (TXOP)[9, 10]. When a sender is allocated a TXOP, it informs surrounding stations (STAs) the time that the channel will remain busy. However, many times the transmission finishes before the reserved time and channel remains idle. In bidirectional method, the receiver STA is allowed to send packets to the sender STA in the reverse direction for the remaining TXOP time. This feature is useful in sending small feedback packets to the sender during the actual data transmission period.

802.11 has been extensively analyzed and various models have been proposed in order to better understand the performance of 802.11 DCF throughput. DCF is a carrier sense multiple access with collision avoidance (CSMA/CA) scheme with binary slotted exponential backoff. Bianchi's analytical model [11] is one of the widely used schemes that is not only simple but it can also predict accurately system throughput for a number of wireless stations in ideal channel conditions. Bianchi's model treats the backoff window size in the protocol as bi-dimensional Markovian chain. Using this chain, Bianchi attempts to compute the probability that a station transmits in a randomly chosen slot time and ultimately derives normalized system throughput as the fraction of time the channel is used to successfully transmit payload bits.

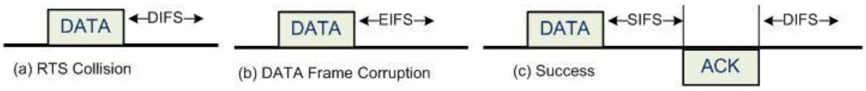
We have used Bianchi's analytical model to analyze various schemes proposed to enhance the saturation throughput in 802.11n. Our contribution is the comprehensive analytical treatment of 802.11n by exploring several enhancement schemes proposed for the latest protocol. Our analysis encompasses the following important scenarios:

1. DCF two-way handshake
2. DCF four-way handshake

Fig. 1: DCF two-way handshake transmission sequence

3. Aggregation with fragment retransmission (AFR)

4. Aggregated-MPDU (A-MPDU)
5. Aggregated-MSDU (A-MSDU)
6. A-MPDU and A-MSDU with bidirectional data transfer



In each of the above scenarios, we investigate the key parameters involved in the equation for saturation throughput and highlight the changes in those parameters for the case of ideal as well as error-prone channels.

2 Throughput Analysis

For the throughput analysis of 802.11n, we adopt Bianchi's model [11] which models the bi-dimensional process {s(t),b(t)} with discrete-time Markov chain. If $W_i = 2^i CW_{min}$ and the maximum contention window $CW_{max} = 2^m CW_{min}$ then s(t) represents the stochastic process for backoff stage (0 ...m) and b(t) represents the stochastic process for the backoff time counter. Using this model, we analyze the probability of successful transmission and the MAC throughput in the various proposals suggested for 802.11n. In the entire analysis, we will assume that there are fixed number of stations in the WLAN and each transmitting station has saturated traffic, that is, each station is working in full load and has attained stable condition. In other words, each transmitting station has always data to send. Also, we assume that regardless of the number of retransmissions the conditional collision probability for each frame is constant and independent.

The mathematical notation used in this paper is summarized in Table 1. We tried to maintain consistency in the symbols instead of a wide variety of notations used in the literature.

If τ represents the stationary probability that a station (STA) transmits in a randomly chosen slot time, then it can be computed as [11]

$$\tau = \frac{2(1-2p)}{(1-2p)(W+1) + pW\left(1-(2p)^m\right)} \quad (1)$$

Here, p is the unsuccessful transmission probability.

2.1 DCF with Two-way Handshake

First we consider basic Distributed Coordinated Function (DCF) which consists of two-way handshake protocol where a station sends a frame and waits for its acknowledgment in a unidirectional channel. The possible time sequence for basic DCF is shown Fig. 1. The normalized throughput (S) for ideal channel and error-prone channel are calculated as follows.

Symbol	Meaning
τ	Stationary probability that a station transmits in randomly chosen time
W	Minimum congestion window size
m	Maximum backoff stage, $i \in (0, m)$ where $W_i = 2^i W$
n	Total number of stations
f	Number of fragments or sub frames in an aggregated frame
$E[P]$	Expected payload
T_e	Virtual time slot length for error transmission sequence
T_s	Average time the channel is sensed busy because of successful transmission
T_c	Average time the channel is sensed busy by non-colliding stations because of a collision
T_{sifs}	Time duration for transmitting a SIFS
T_{difs}	Time duration for transmitting a DIFS
T_{eifs}	Time duration for transmitting an EIFS
T_{ack}	Time duration for transmitting an ACK
$T_{phy\ hdr}$	Time duration for transmitting a physical header
$T_{mac\ hdr}$	Time duration for transmitting a MAC header
T_f	Time duration for transmitting one AFR frame payload
L_{frag}	Fragment length in bytes
L_{fcs}	FCS length in bytes
L	Total MAC frame length in bytes
L_{hdr}	Total length of MAC header in bytes
$L_{hdr+fcs}$	Total length of MAC header and FCS in bytes
L_{data}	Full payload of MAC frame in bytes
p	Probability of unsuccessful transmission
p_b	Probability of single bit error
p_e	Error probability for non-ideal channel
p_c	Collision probability

Table 1: Mathematical Notation Ideal Channel For an ideal channel, the successful transmission means only one STA transmits out of n STAs at any given time. The probability that a STA does not transmit in randomly chosen slot time is $1 - \tau$. The probability that $n - 1$ STAs do not transmit will be $(1 - \tau)^{n-1}$. For an ideal channel, the unsuccessful transmission is because of collision and hence the probability that a transmitted packet encounters a collision (p) is given by

$$p = pc = 1 - (1 - \tau)^{n-1} \quad (2)$$

Since the probability that no STA transmits is $(1 - \tau)^n$, the probability Ptr that at least one station transmits will be

$$Ptr = 1 - (1 - \tau)^n \quad (3)$$

The probability of exactly one transmission is $\binom{n}{1}\tau(1 - \tau)^{n-1}$. Therefore, the probability Ps that a transmission is successful is given by the probability that exactly one station transmits, condition on at least one station transmits [11]

$$Ps = \frac{n\tau (1 - \tau)^{n-1}}{1 - (1 - \tau)^n} \quad (4)$$

The normalized throughput S is the ratio of expected payload transmitted in a slot time to the expected length of a slot time. If E[P] is average packet payload size, then

$$S = \frac{PtrPsE[P]}{P1\sigma + P2Ts + P3Tc} \quad (5)$$

Where P1, P2 and P3 are the probabilities that a slot is empty, it contains a successful transmission, and it contains a collision respectively and σ is the duration of an empty slot time. Thus, $P1 = (1 - Ptr)$, $P2 = PtrPs$, $P3 = Ptr(1 - Ps)$. Ignoring the transmission delay, Ts and Tc can be written as (see Fig. 1):

$$Ts = Tdata + Tsifs + Tack + Tdifs \quad (6)$$

$$Tc = Tdata + Teifs \quad (7)$$

The expected payload for DCF two-way handshake under ideal channel conditions is given by: $E[P] = (L - Lhdr + fcs)$ (8)

Channel with Error Since the channel is error prone, the unsuccessful transmission could be due to frame collision or channel error. Hence, the unsuccessful transmission probability (p) in Eq. 1 needs to be adjusted for both collisions and transmission errors. Thus, p can be expressed as

$$p = 1 - (1 - pc)(1 - pe) \quad (9)$$

where $pc = 1 - (1 - \tau)^{n-1}$ is the collision probability and pe is the error probability on the condition that there is a successful transmission in the time slot.

The throughput Eq. 5 can now be written for error prone channel as

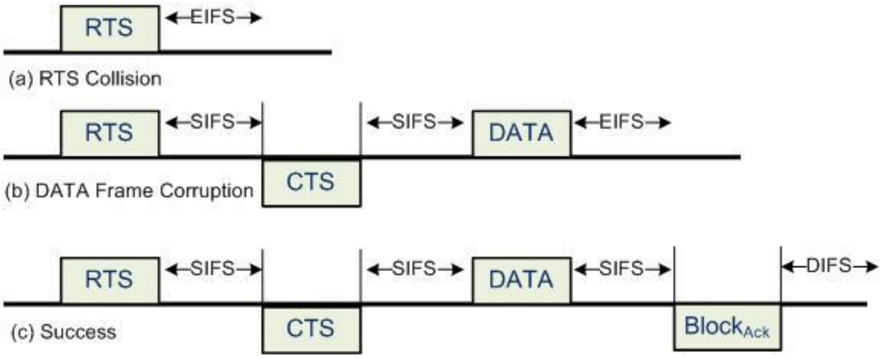

RTS
EIFS
(a) RTS Collision
RTS
SIFS
CTS
SIFS
DATA
EIFS
(b) DATA Frame Corruption
RTS
SIFS
CTS
SIFS
DATA
SIFS
Block$_{Ack}$
DIFS
(c) Success

Fig. 2: RTS/CTS unidirectional access

$PtrPsE[P\ ]$

$S = (10)$

$P1\sigma + P2Ts + P3Tc + peTe$ where Te is the virtual time slot length for error transmission sequence. It can be given as $Te = Tdata + Teifs$ (11)

Tc and Ts remain the same. Also note that in Eq. 10, P2 is now equal to $PsPtr(1 - pe)$.

The only unknowns left out in Eq. 10 are pe and E[P]. As we are aware that any single bit error would corrupt the entire frame, thus

$$pe = 1 - (1 - pb)^{L} \quad (12)$$

$$E[P\ ] = (L - Lhdr+fcs)(1 - pe) \quad (13)$$

2.2 DCF with Four-way Handshake

DCF with four-way handshake uses RTS/CTS (Ready to Send/Clear to Send) control packets. A station that wishes to send a data frame exchanges RTS/CTS frames with the intended recipient before sending the actual data frame. This mode fixes the “hidden node” problem. The possible time sequence for DCF with four-way handshake in unidirectional data transfer is shown in Fig. 2. DCF four-way throughput analysis is similar to DCF two-way analysis of Section 2.1. The normalized throughput is given by Eq. 5 and Eq. 10 for ideal and error-prone channel conditions respectively. The only difference is in the duration of Ts, Tc and Te, which are given by:

$$Ts = Trts + 3Tsifs + Tcts + Tdata + Tack + Tdifs \quad (14)$$

$$Tc = Trts + Teifs \quad (15)$$

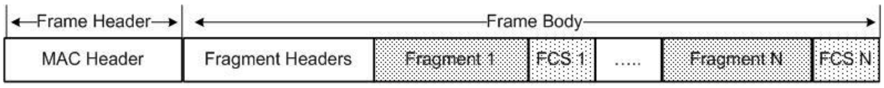


Fig. 3: Aggregation with fragment retransmission

Te = Trts + Tcts + Tdata + Teifs +2Tsifs (16)

2.3 Aggregation with Fragment Retransmission (AFR)

AFR [4] scheme works by combining several packets into one large frame and dividing this frame into multiple fragments. Fragmentation is done to avoid retransmission of the whole frame in case of any packet corruption during the transmission. Only the fragment containing the corrupted packet is retransmitted, thus increasing the overall performance. AFR also employs zero-waiting policy in which frames don't wait for all the fragments from upper layer to arrive. Instead, frames are transmitted as soon as a station wins transmission opportunity. Thus, AFR offers improved throughput through aggregation, fragmentation for selective retransmission and zero-waiting scheme.

AFR can also be modeled using Bianchi's model and the throughput Eq. 5 can be directly applied for ideal as well as error-prone channels as follows:

$$Safr = \frac{P2E[L]}{P1\sigma + P2Ts + P3Tc} \quad (17)$$

It should be noted that AFR considers partially corrupted frames due to channel noise as successful transmission. Since E[L] is the expected number of successfully transmitted bits instead of frame payload size, it can be calculated as:

f��

f

frag frag

$$E[L]= \cdot (p)^{i} \cdot (1 - p_e)^{f-i} \cdot (L - i \cdot Lfrag)$$

e

i

i=0

frag

where fragment error rate pis given as

e

frag $=1 - (1 - pb)$Lfrag+Lfcs

p(18)

e

Lfcs is added in Eq. 18 as each fragment in AFR data frame has FCS (see Fig. frag
3). Substituting the value of pin Eq. 17 and simplifying to get

e frag

$\underline{P2 \cdot L \cdot (1 - p)}$

$\underline{e}$

$Safr = (19)P1\sigma + P2Ts + P3Tc$

The other unknown values Ts and Tc in Eq. 17 are different for DCF two-way handshake and DCF four-way handshake.

In case of DCF two-way handshake, Ts and Tc are given as:

afr

$$Ts = T_{data} + Tsifs + Tack + Tdifs \quad (20)$$

$$T_c = T^{afr}_{data} + T_{eifs} \quad (21)$$

where $T^{afr}_{data} = T^{phy}_{hdr} + T_f$.

For DCF four-way handshake, Ts and Tc are given as:

$$T_s = T_{rts} + T_{cts} + T^{afr}_{data} + 3T_{sifs} + T_{ack} + T_{difs} \quad (22)$$

$$T_c = T_{rts} + T_{eifs} \quad (23)$$

1.4 A-MPDU/A-MSDU Frame Aggregation

In this section, we will model A-MPDU and A-MSDU schemes using Bianchi’s model [11]. Aggregated A-MPDU and A-MSDU frames can be transmitted using either DCF two-way handshake or DCF four-way handshake, and hence the analysis of A-MPDU and A-MSDU for DCF two-way hand shake and four-way hand shake will be similar to the standard analysis of Sections 2.1 and 2.2 respectively. The normalized throughput for A-MPDU and A-MSDU is given by Eqs. 5 and 10 respectively for ideal and error-prone channel conditions. However, the expected payload values (E[P]) in these equations depend upon the aggregation method and channel condition, and can be estimated as follows.

Ideal Channel A-MSDU and A-MPDU frame structure is shown in Fig. 4. Assume that there are f subframes in the each aggregated A-MSDU and AMPDU frame.

In A-MSDU, for each subframe there is an additional overhead of subframe header (14 bytes) and padding (0-3 bytes). Hence, the expected payload size for A-MSDU is

$$E[P] = L - L^{oh}_{a-msdu} \quad (24)$$

where $L^{oh}_{a-msdu} = L_{hdr+fcs} + \sum_{i=1}^{f}(L_{subhdr} + L_{pad})$.

On the other hand, each subframe in A-MPDU has a separate MAC header, a delimiter (4 bytes), variable size padding (0-3 bytes) and FCS. Hence the expected payload for A-MPDU is given by:

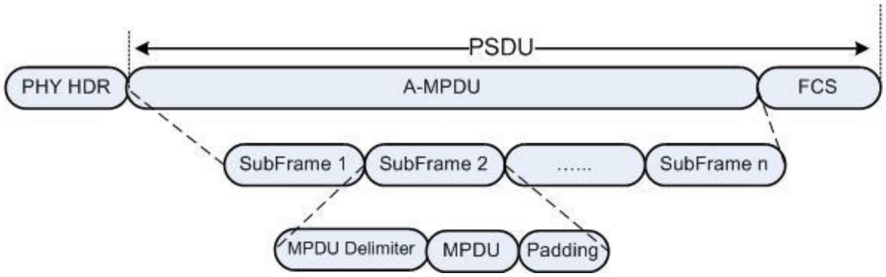
$$E[P] = L - \sum_{i=1}^{f}(L_{hdr+fcs} + L_{dlim} + L_{pad}) \quad (25)$$

The above equation can be rewritten as
f
$(L_i - L^{oh}$

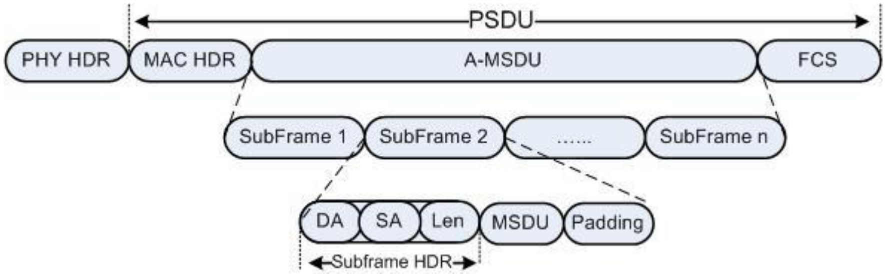
$E[P\ ]= {}_{a-mpdu})$ (26) i=1

where Li is ith subframe of the aggregated A-MPDU frame and L^{oh} = a−mpdu Lhdr+fcs + Ldlim + Lpad. It should be noted that Ts is slightly different for A-MPDU/A-MSDU as single block ACK will be sent instead of individual ACKs.

Ts = Trts + Tcts + Tdata + Tback +3Tsifs + Tdifs (27)



(a) A-PMDU frame structure



(b) A-MSDU frame structure

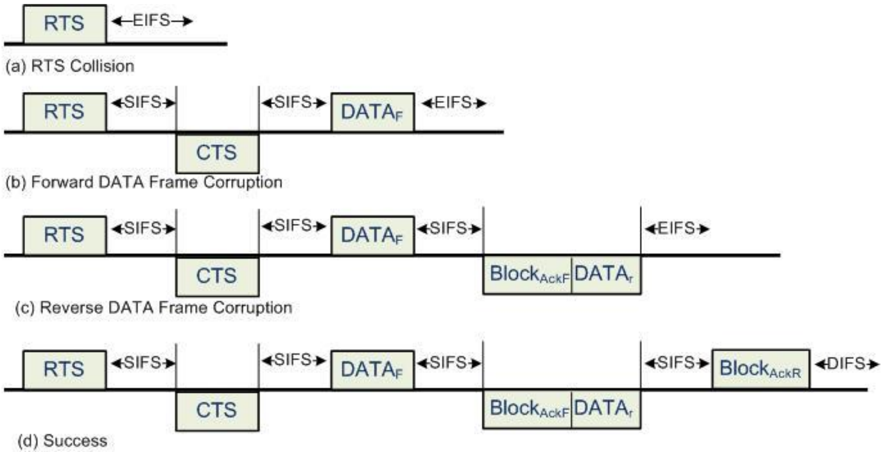
Fig. 4: One-level frame aggregation

Error Prone Channel In order to calculate the values for pe and E[P] in Eq. 10, we first consider the case for A-MSDU where any single bit error would corrupt the entire frame, thus $pe = 1 - (1 - pb)^L$ (28) $E[P\ ]=(L - L^{oh}_{a-msdu})(1 - pe)$ (29) Fig. 5: RTS/CTS bidirectional data transfer

In the case of A-MPDU, error occurs when all the sub-frames are corrupted. Therefore, we can write

where i ranges from 1 to the total number of sub-aggregated MPDUs frames (f).

2.5 Bidirectional Data Transfer



pe = Π (1 − (1 − pb)Li) (30)

i=1

E[P] = f i=1 (Li − Loh a−mpdu)(1 − pb)Li (31)

A key enhancement in 802.11n specifications is bidirectional data transfer. In this section, we will extend the A-MSDU/A-MPDU analysis of Section 2.4 for bidirectional data transfer. The possible time timing sequence for DCF four-way handshake with bidirectional data transfer is shown in Fig. 5. The DATA frames in this figure represent A-MPDU/A-MSDU aggregated frames.

The normalized throughput equations for bidirectional data transfer remain the same and are given by Eqs. 5 and 10 respectively for ideal and error-prone channel conditions. However, the expected payload value (E[P]) should accommodate both forward and reverse payloads. In the ideal channel case, E[P] is the sum of forward and reverse payloads. If we assume that the data in the reverse direction is also aggregated and it contains b subframes then E[P] for A-MSDU and A-MPDU is given by the following equations respectively.

E[P]= Lf + Lb − 2 ∗ L^{oh} (32)

a−msdu fb $_{(L_i - L}$oh $_{(L_i - L}$oh

E[P] = (33)

a−mpdu$^{)+}$ a−mpdu$^{)}$ i=1 i=1

Also Ts should be modified to accommodate the reverse payloads, and it is given by:

ff

Ts = Trts +4Tsifs + Tcts + T + T + T ^{r} (34)

data $^{+ T^{r}}$

data back back $^{+ T}$difs

Error Prone Channel In this case, an error can occur during the forward or reverse transmission. Hence the error probability has two components (pe,f , pe,r) corresponding to figure 5(b) & 5(c). The virtual time slots corresponding to error transmission in the forward and reverse directions are given by:

f
Te,f = Trts +2Tsifs + Tcts + T + Teifs (35)
data
ff $_{+\,T}$ r
Te,b = Trts +3Tsifs + Tcts + T + T $_{data}$ + Teifs (36)
data back

The error probabilities in the forward and reverse directions are dependent on aggregation method.

For A-MSDU, the error probabilities in the forward sequence and reverse sequence are given by

$$p_{e,f} = 1 - (1 - P_b)^{L_{for}} \quad (37)$$

$$p_{e,r} = (1 - P_b)^{L_{for}} (1 - (1 - P_b)^{L_{rev}}) \quad (38)$$

The expected payload for A-MSDU in the bidirectional mode is:

$$E[P] = (E_p^f + E_p^r)(1 - p_{e,f} - p_{e,r}) + E_p^f\, p_{e,r} \quad (39)$$

where E_p^f and E_p^r are the expected payloads given by Eq. 24.

In case of A-MPDU, the error probability in the forward direction can be calculated as:

f
$_{(1-(1-p_b)}$Li,f $_{)}$
pe,f = (40) i

The error probability in the reverse direction (pe,r) is dependent upon the success of the forward transmission. Assuming the reverse data is also aggregated:

f
$_{(1-(1-p_b)}$Li,r $_{)]}$
pe,r = (1 − pe,f)[(41)
i

The expected payload for A-MPDU in bidirectional transmission is given by Eq. 39, where E_p^f and E_p^r are the expected payloads in the forward and reverse directions. These can be calculated by using Eq. 26.

2.6 Conclusion and Future work

In this paper we analyzed and compared the normalized MAC throughput for various aggregation schemes in 802.11n using Bianchi's analytical model. The following are some conclusions that can be drawn from the analysis.

In ideal channel conditions A-MSDU performs very well since there is no overhead of MAC headers & FCSs. Conversely, when BER increases A-MPDU outperforms A-MSDU where single bit error corrupts the whole A-MSDU aggregated frame. A-MPDU obviates the need to resend the entire aggregated frame since the receiver can delineate a received A-MPDU frame and sends a BlockAck allowing individual data frames to be acknowledged or retransmitted.

Bidirectional data transfer provide significant improvement over unidirectional data transfer when receiver has always data to send in the reverse direction (for example, applications like voice chatting). On the other hand, it won't add any advantage over unidirectional data transfer in terms of MAC throughput if the data is predominantly unidirectional, except that higher layer protocol can benefit from reverse data in terms of timely acknowledgments.

Larger frame size increases the probability of collision (P3) thereby decreasing the throughput. Since aggregation schemes employ larger frame size, they can benefit from four way handshake (RTS/CTS) to reduce the probability of collision because of smaller RTS/CTS control frames. AFR scheme reduces the probability of error (P^{frag}) by fragmenting the frame and selectively retrans emitting the erroneous fragment thus improving the overall throughput.

In future, we would like to simulate aggregation schemes and compare the simulation results with our analysis. We would also like to analyze and simulate multi-level aggregation schemes and compare them with the existing schemes.

References

1. IEEE Computer Society, “Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Enhancements for Higher Throughput”, IEEE P802.11n-D2.0, Feb 2007.
2. K. Feng, Y. Huang, J. Lin “Design of MAC-defined aggregated ARQ schemes for IEEE 802.11n networks”, Wireless Networks 17(3), pages: 685-699, 2011.
3. Y. Xiao, “IEEE 802.11n: Enhancements for higher throughput in wireless LANs,” IEEE Trans. Wireless Communication, 12 (6), pages: 8291, Dec 2005.
4. T. Li, Q. Ni, D.Malone, D. Leith , Y. Xiao and T. Turletti, “Aggregation with fragment retransmission for very high-speed WLANs,” IEEE/ACM Transactions on Networking, 17 (2), pages: 591-604, Apr 2009.
5. J. Tourrilhes, “Packet frame grouping: Improving IP multimedia performance over CSMA/CA,” in Proc. ICUPC, pages: 13451349, 1998
6. V. Vitsas, P. Chatzimisios, A. C. Boucouvalas, P. Raptis, K. Paparrizos, and D. Kleftouris, “Enhancing performance of the IEEE 802.11 distributed coordination function via packet bursting,” IEEE GLOBECOM, 2004.
7. Y. Xiao and J. Rosdahl, “Performance analysis and enhancement for the current and future IEEE 802.11 MAC protocols,” ACM SIGMOBILE Mobile Computing and Communications Review, vol. 7, no. 2, pages. 619, Apr 2003.
8. Y. Lin and V.W.S. Wong, “Frame Aggregation and Optimal Frame Size Adaptation for IEEE 802.11n WLANs”, IEEE Globecom, Nov 2006.
9. C. Liu and A. P. Stephens, “An analytic model for infrastructure WLAN capacity with bidirectional frame aggregation,” in Proc. of IEEE WCNC, pages: 113119, Mar. 2005
10. D. Skordoulis, Q. Ni, H. Chen, A. Stephens, C. Liu, and A. Jamalipour, “IEEE 802.11n MAC frame aggregation mechanisms for next-generation high-throughput WLANs,” IEEE Wireless Communications, pages: 40-47, Feb 2008.
11. G. Bianchi, “Performance analysis of the IEEE 802.11 distributed coordination function”, IEEE Journal of Select. Areas Communication, vol. 18, no. 3, pp. 535547, Mar. 2000.