

Investigation of Techniques for VoIP Frame Aggregation Over A-MPDU 802.11n

Qasem M. Kharma*, Abdelrahman H. Hussein, Faris M. Taweel, Mosleh M. Abualhaj and Qusai Y. Shambour

Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan

*Corresponding Author: Qasem M. Kharma. Email: q.kharma@ammanu.edu.jo

Received: 23 May 2021; Accepted: 24 June 2021

Abstract: The widespread and desirable features of IP and IEEE 802.11 networks have made these technologies a suitable medium for carrying voice over IP (VoIP). However, a bandwidth (BW) exploitation obstacle emerges when 802.11 networks are used to carry VoIP traffic. This BW exploitation obstacle is caused by the large 80-byte preamble size of the VoIP packet and a waiting time of 765 μ s for each layer 2 VoIP frame. As a solution, IEEE 802.11n was consequently designed with a built-in layer 2 frame aggregation feature, but the adverse impact on the VoIP performance still needed to be addressed. Subsequent VoIP research has proposed numerous enhancements to the built-in IEEE 802.11n aggregation techniques. In this study, we analyse and propose enhancements to the BW exploitation and VoIP performance over IEEE 802.11n. We also recommend new techniques to enhance VoIP performance via IEEE 802.11n frame aggregation implementations. Furthermore, we contribute guidelines for creating a robust, suitable, and reliable aggregation technique for VoIP over the 802.11n standard.

Keywords: A-MPDU; A-MSDU; VoIP; bandwidth exploitation; VoIP frame aggregation; IEEE 802.11n

1 Introduction

In recent years, the expansion of voice over IP (VoIP) applications, such as Zoom and Facebook Messenger, has become progressively omnipresent. More than 1.3 million clients who had been active on Facebook Messenger were recorded monthly in 2018. In addition, the Internet has consistently moved in excess of 156 petabytes of VoIP information since 2015 [1,2]. The enormous universality of VoIP has numerous motivators, in which the most significant is the faint or free call cost. Another notable incentive is the ability to conjugate voice with other useful features, including audio conferencing, text chatting, and video conversation. Another example of the critical incentive is the portability of VoIP, i.e., people can make calls by using most types of modern devices (smart phone, tablets, and laptops) [1,3].

The IEEE 802.11 standard, which manages the implementation of wireless local area networks, has also fundamentally proliferated in all sectors owing to its installation simplicity [4,5]. Furthermore, the IEEE 802.11 standard has received numerous enhancements, which has subsequently led to IEEE



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

802.11 standard enhancements [5,6]. In the previous 802.11 standards, the physical (PHY) and medium access control (MAC) layers had experienced several upgrades, including the introduction of multiple input multiple output (MIMO) technology at the PHY layer, and the doubling of the channel bandwidth (BW) to 40 MHz with the orthogonal frequency division aggregation (MIMO-OFDM) and the MIMO technology. Consequently, the data rate of this layer has been improved by up to 600 Mbps. However, the MAC layer also requires to be updated to ensure that the upgrades at the PHY layer are effective. The principal upgrades in the MAC layer are the block acknowledgement (BA) and frame aggregation [4,7,8]. This study focuses on frame aggregation. Addressing the problem of VoIP and 802.11 alongside the upgrades suggested in the .11n networks may require an integration of all areas to be able to send voice data over 802.11 networks [5,9].

Two main issues arise with the aggregation of voice data over 802.11 networks. First, the quality of VoIP calls is debased because of the rise in packet losses, delays, and jitters (variations of delay) [1,10,11]. Second, the large preamble size of the 802.11 frame prompts a squandering of the available BW [12–14]. This study concentrates more on this large preamble size and less on the quality issue. The large preamble size issue has been addressed by numerous approaches. One of these techniques is the VoIP packet preamble compression. Preamble compression techniques have succeeded in compressing the 40-byte VoIP packet preamble into 2–4 bytes. Another key approach is VoIP packet aggregation. The packet aggregation approach consolidates numerous frames into a single large frame. The 802.11n amendment has proposed two aggregation techniques as part of its design, namely, the aggregation MAC service data unit (A-MSDU) and the aggregation MAC protocol data unit (A-MPDU) [13,15,16]. Ref. [16] has investigated A-MSDU techniques. In this research, we investigate the A-MPDU packet aggregation techniques of VoIP packets over the .11n standard. This study particularly investigates the packet aggregation approach of VoIP packets over the .11n network. Our work also explores and underscores the wastage of the BW when running VoIP over 802.11n networks.

The paper is composed of six sections. The introduction, including the main aim of this paper, is presented in Section 1. The key topics related to this work are presented in Section 2. The current VoIP aggregation techniques over 802.11n networks are presented and analyzed in Section 3. An investigation and the analysis of frame aggregation on BW exploitation are presented in Section 4. Section 5 provides a set of guidelines on how to design a new aggregation technique. Section 6 presents the conclusion.

2 Background

The two key topics related to this work are discussed in this section. The first topic is on VoIP over IEEE802.11n, and the second topic is about VoIP packet aggregation techniques.

2.1 VoIP Over 802.11n

VoIP service progressively rules the media communications sector by permitting voice calls to be carried as digital network packets. VoIP fineness is profoundly influenced (degraded) when packet losses and delays are increased [17]. The most elevated portion of the delay is caused by a codec creating the voice frame (VoIP packet payload). In certain cases, the VoIP packet payload consists of more than one voice frame. Tab. 1 presents a number of voice codecs, in which the average frame size can reach 30 bytes [18,19]. For each VoIP packet frame, an 80-byte RTP/UDP/IP/Ethernet preamble (12-byte real-time transport protocol, 40 Ethernet 802.11n, 20-byte Internet protocol, and 8-byte user datagram protocol) is appended to form the VoIP layer 2 frame when using the 802.11n standard [6,20,21]. The 802.11n frame format is presented in Fig. 1.

The maximum speed of 600 Mb/s in the 802.11n is deliberated. The time needed to send the 80-byte preamble is $80 \times 8/600 = 1.06 \mu\text{s}$. Furthermore, the VoIP packet needs 765 μs , distributed among backoff

(BO) time, short inter-frame space (SIFS), distributed inter-frame space (DIFS), and ACK, prior sending. The subsequent expended time is approximately equivalent to 766 μs (765 and 1.06 μs). The required time to send a 20-byte voice frame by using G.726 is $20 \times 8/600 = 0.27 \mu s$. Consequently, the voice frame portion from the BW is insignificant and equivalent to roughly 0.000261 (0.27 μs divided by 766 μs). The proportion of shifts is reliant on the length of the voice frame. Unfortunately, this method causes extraordinary wastage of the available BW [1,22,23]. Fig. 2 illustrates the squandered BW when a layer 2 VoIP frame is transferred at a given 20-byte G.729.

Table 1: Examples of VoIP codecs

Codec	G.729	G.728	G.723.1	G.726	LPC
Frame size (Byte)	10	10	20/24	15	14
Frame size (ms)	10	5	30	5	20

2 bytes	2 bytes	6 bytes	6 bytes	6 bytes	2 bytes	6 bytes	2 bytes	4 bytes	0.7955 bytes	4 bytes
Frame Control	Duration /ID	Address 1	Address 2	Address 3	Seq. Control	Address 4	QoS Control	HT Control	Body	FCS

Figure 1: Frame format in 802.11n

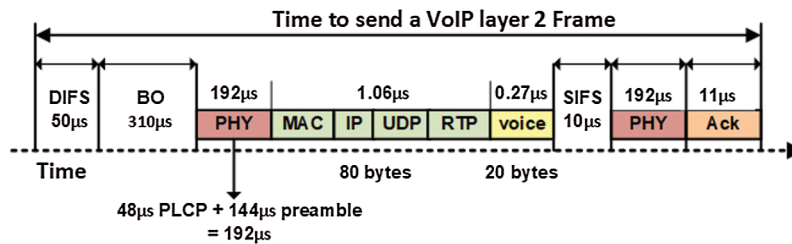


Figure 2: Time to transmit a layer2 VoIP frame

A noticeable enhancement of the .11n standard is A-MPDU frame aggregation at layer 2. This layer 2 frame aggregation gathers numerous frames in an enormous A-MPDU by using a single preamble only. Thus, the proportion of the layer 2 frame payload to the preamble size is expanded. Moreover, the holding-up timing spans among numerous frames are decreased because a single enormous A-MPDU is sent rather than many layer 2 frames with a single-channel access. On this basis, the BW use effectiveness can be exceptionally improved, and the limit of the synchronous calls can be expanded subsequently. The general process of layer 2 frame aggregation is as follows. First, the aggregating entity at the transmitter side inspects the frame receiver, groups the frames traveling in the same path into a single preamble, and generates and transmits the A-MPDU to the receiver. When the A-MPDU reaches the receiver side, the de-aggregating entity breaks the A-MPDU and restores the grouped layer 2 frames. Fig. 3 illustrates the general frame aggregating/de-aggregating process. The aggregation of three 802.11n frames into a single frame is illustrated in Fig. 4. As shown by the figure's upper segment, the time needed to independently send three frames is 2298 μs (766 * 3). Meanwhile, as shown in the figure's lower segment, the time needed to send an aggregated frame produced from the aggregation of three frames is virtually 766 μs. In this manner, the frame aggregation techniques consume less time to transmit the data [12,24–27].

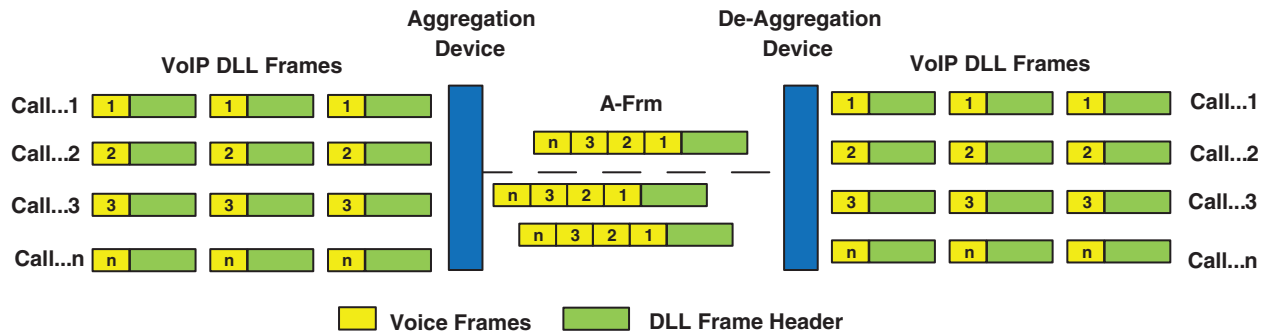


Figure 3: Aggregating/De-aggregating process

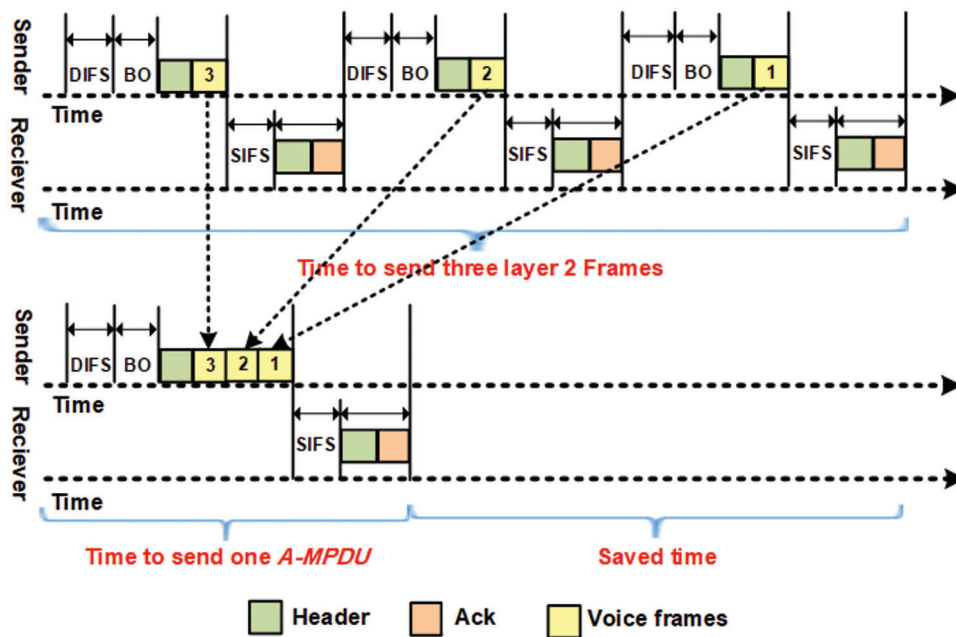


Figure 4: Reducing time intervals and preamble overhead via frame aggregation

2.2 A-MPDU IEEE 802.11n Aggregation Technique

The wireless layer 2 frame is technically called MPDU. The MPDU includes a trailer, a body (MSDU), and a preamble. The A-MPDU operates at the bottom part. Several MPDUs sent to a single receiver are joined in a single A-MPDU with one 802.11n PHY preamble. Then, the A-MPDU is transferred to the PHY layer as a single PHY layer service data unit (PSDU) and prepared as a single PHY layer protocol unit (PPDU) to be transmitted to the channel. Each sub-frame in the A-MPDU consists of the sub-frame MPDU delimiter (4B), MPDU frame, and padding (0–3B). The delimiter consists of the reserved (4-bit), MPDU length (12-bit), CRC (8-bit), and delimiter (8-bit) signatures. Each MPDU frame consists of a MAC preamble and an MSDU. Fig. 5 depicts the A-MPDU frame structure. Similar to the A-MSDU, all MPDUs within a single A-MPDU must have only a single transmitter and receiver address. However, this address is not essentially the same destination and source address. Moreover, in cases in which the video, voice, and best-effort data cannot be grouped, the same QoS category must be implemented in all sub-frames within the A-MPDU. With the A-MPDU, the BA can then be used, in which each aggregated MPDU is ACKed separately. Therefore, different from the A-MSDU, if any of the aggregated MPDUs is

damaged, then it can be retransmitted by itself without the need to retransmit the entire A-MPDU. Thus, the A-MSDU is more efficient than the A-MPDU because the former is encapsulated with a single 802.11n MAC preamble. However, the A-MPDU is more efficient than the A-MSDU in environments with high BER and FER rates because only the damaged MPDUs need to be retransmitted in the A-MPDU technique. The entire A-MSDU incidentally needs to be retransmitted in the A-MSDU technique if an error occurs. Moreover, the A-MPDU is formed from the frames within the buffer and does not wait for more frames, suggesting that the A-MPDU can be formed immediately. Tab. 2 summarizes the key differences between the A-MSDU and A-MPDU aggregation techniques [22,28–31].

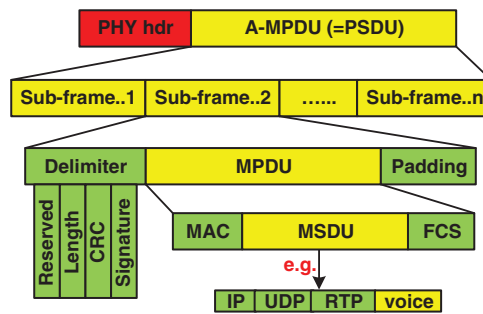


Figure 5: A-MPDU frame structure

Table 2: Key differences between A-MSDU and A-MPDU

Attribute	A-MSDU	A-MPDU
MAC preamble overhead	Lower	Higher
Complexity	Lower	Higher
Aggregation location	Top part of layer	Bottom part of layer
Multiple destination	No	Yes
Multiple receiver	No	No
Max length (B)	3839 or 7935	8191 or 16383 or 72767 or 65535
Max no. of sub-frames	No limit	64
Sub-frames retransmission	Not possible	Possible
Encryption	Encrypt the entire A-MSDU in one operation	Encrypt each sub-frame separately
Traffic class	One for all sub-frames	One for each sub-frame

Apart from the IEEE 802.11n standard aggregation techniques (A-MSDU and A-MPDU), two other aggregation techniques are proposed at the PHY layer, namely, the aggregation PHY service data unit (A-PSDU) and the aggregation PHY protocol data unit (A-PPDU). The A-PSDU is used for the top section of the PHY layer where many PSDUs are combined. The A-PSDU distinguishes the aggregated PSDUs by adding a PHY signaling field to denote the PSDU duration. The modulation coding scheme (MCS) for the 802.11n standard does not consider the A-PSDU given the robustness of the A-MPDU in the

combining frames that travel to the same destination. Moreover, the A-PSDU adds a high overhead. Any error, even if only 1 bit is involved, hinders the demodulation of the remaining frame. However, the A-PPDU (PPDU burst) resides at the bottom of the PHY layer. In this technique, several frames originating from a single source and with one or more destinations for the same transmission power level are combined. The PPDU is distinguished using inter-frame spacing. In addition to these aggregation techniques, many other studies have proposed to combine the A-MSDU and A-MPDU techniques into a single technique called multi-level/two-level frame aggregation. At the first level, several MSDUs are assembled for the A-MSDU construction. At the second level, several A-MSDUs are grouped into a single A-MPDU, in which the entire A-MSDU frame becomes a single sub-frame in the A-MPDU aggregation. The two-level aggregation attempts to overcome the shortcomings of the A-MSDU and A-MPDU techniques, and it can increase the throughput. However, the resulting structure is extremely complex, similarly complicating the aggregation and de-aggregation processes [22,28–31]. Fig. 6 presents the frame structure of the two-level aggregation technique.

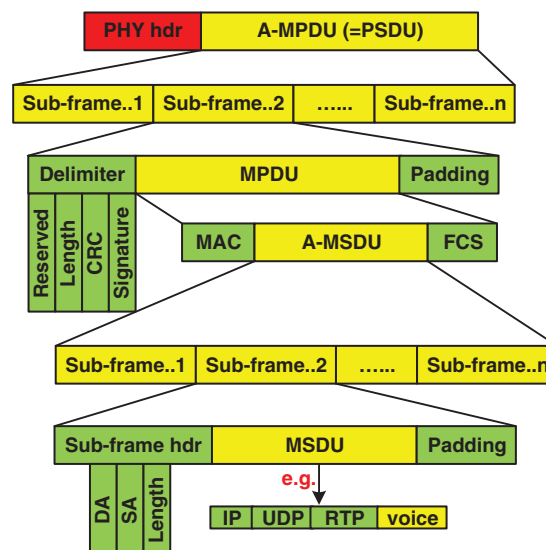


Figure 6: Two-level frame structure

The VoIP service quality requirements (maximum delay, jitter, and packet loss) have not been examined along with the development of the A-MPDU technique. Therefore, the performance of VoIP over A-MPDU 802.11n networks may be seriously retrograded if ordinary traffic is combined with VoIP traffic. Consequently, numerous aggregation techniques have been proposed to deal with the combination of the A-MPDU and VoIP service traffic [14,32,33]. These aggregation techniques are discussed in the following section.

3 VoIP Aggregation Techniques Over A-MPDU 802.11n

The aggregation techniques can be categorized as non-adaptive and adaptive. In the adaptive aggregation techniques, the A-MPDU size is specified and changed dynamically based on the current connection conditions. These conditions include the frame-error rate (FER), signal-to-noise ratio (SNR), congestion, bit error rate (BER), maximum delay, and load. In the non-adaptive aggregation techniques, the A-MPDU size is specified based on preset values of certain parameters, including size, number of packets, and time period. Although each aggregation technique can differently determine the A-MPDU

size, both of their core aggregation processes are similar [12,34,35]. Fig. 7 shows the general core process of the aggregation techniques.

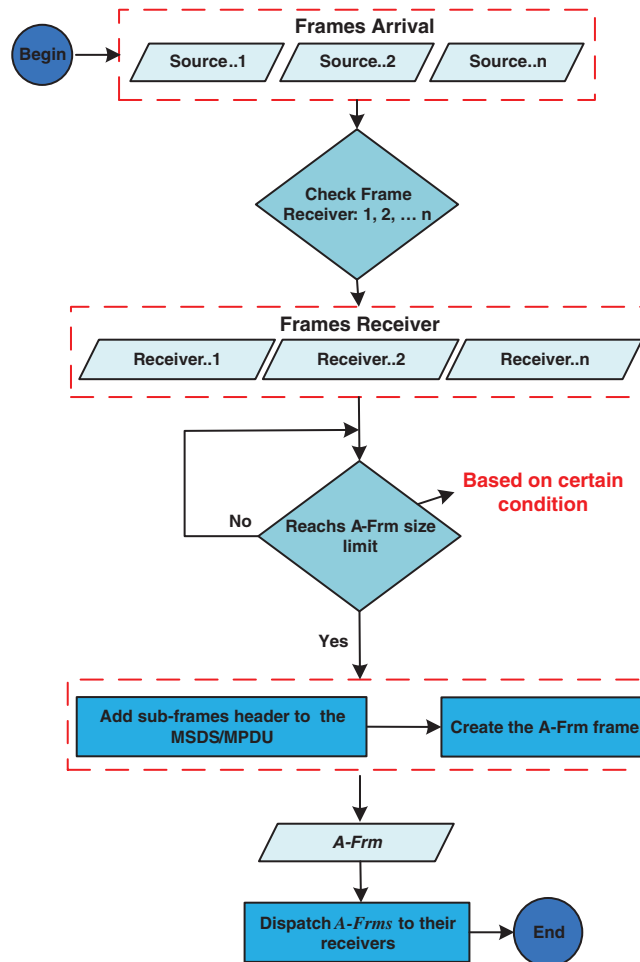


Figure 7: General aggregation process

Ref. [36] proposed a non-adaptive frame aggregation technique that operates at the bottom layers of the MAC. The frames sharing a single route with the same destination are constructed as a massive A-MPDU. Ref. [36] focused on the differences in service of the IEEE 802.11e standard and the arrangement of packets into four ACs, in which the A-MPDU frame length limit was preset. Ref. [36] was able to enhance the IEEE 802.11e and avoid starvation at the low-priority ACs by allowing the aggregation of packets from different ACs into a single A-MPDU. When the predefined size threshold is reached by the length of the frame in the buffer, the aggregation process is stopped, and the A-MPDU frame is transmitted to the receiver. Similar to the configuration of the traditional A-MPDU, the entire A-MPDU is not retransmitted, only the damaged MPDUs. The implementation of this technique is illustrative of an assessment of delay improvement. In the implementation of the testing scenario, the delay could be decreased in terms of VoIP, and the VoIP quality was consequently enhanced.

Different from the pre-discussed technique, Refs. [14,33,37,38] proposed the use of the adaptive frame aggregation technique. This proposed technique operates at the lower part of the MAC. The technique groups the frames intended for the same receiver into a single A-MPDU frame. In the conventional A-MPDU, when

any of the aggregated MPDUs is damaged, only the retransmission of the MPDU is triggered. Besides, the proposed technique adopts the variation of service in IEEE 802.11e networks and clusters the packets into four ACs, each with a preset A-MPDU frame length limit based on the AC requirements. However, although the technique can handle the four ACs, its main focus target is the voice AC. The proposed scheduler uses the core merits of the ath9k driver with some added merits, including the transmission queue (TxQueue), software queue (SwQueue), new features in SwQueue, new features in TxQueue, and A-MPDU size restrictions. The A-MPDU maximum size changes adaptively on the basis of the average medium access delay statistics and the time-varying maximum end-to-end delay acquired from the RTP/RTCP protocols, tolerable delay (150 ms in the case of VoIP), predefined buffering delay thresholds, and channel congestion. The performance analysis of the proposed technique showed excellent throughput improvement while maintaining the delay and packet loss ratio within the tolerable QoS range. However, the implementation of the proposed technique in an ath9k driver of an actual 802.11 Wi-Fi testbed with a PHY transmission rate of 150 Mbps in 802.11n obtained an improvement of 160% compared with the default driver implementation. Additional experiments were subsequently conducted on the NS-3 simulator, in which 50 nodes were transmitted at the PHY transmission rate of 600 Mbps. The default configuration provided an approximate of 0.13 Mbps throughput only with only 4% of voice packet delivery. The proposed technique could provide up to 275 Mbps of throughput with 100% voice packet delivery and an end-to-end delay of less than 150 ms.

Similar to the previous techniques [14,33,37,38], Refs. [32,39] proposed the adaptive technique that operates at the bottom part of the MAC layer. The frames intended for a single receiver are grouped into a single A-MPDU frame. When any of the aggregated MPDUs is damaged, the MPDU alone is retransmitted. Besides, the frames in the proposed technique were divided into three categories, namely, voice, video, and streaming categories with priority queuing (PQ), which assume a scheduling. The A-MPDU maximum size changes adaptively on the basis of urgency delay (UD), which represents the remaining time to dequeue the frame. The frame with the lowest UD is the first packet added to the A-MPDU. On the basis of the remaining UDs of this frame, the maximum aggregation time of the A-MPDU is subsequently set. In addition to frame aggregation, a new scheduling scheme was proposed to overcome the drawbacks of the PQ scheduler under saturated conditions. The performance analysis of the proposed technique results showed a higher throughput rate and lower delay and packet loss compared with those of the other comparable techniques.

Ref. [40] proposed a distinct technique that deploys two-level frame aggregation. In the first level, the MSDUs with the same recipient are grouped into an A-MSDU. In the second level, the A-MSDUs with the same recipient are grouped into an A-MPDU. Similar to the common A-MPDU, the damage of any of the MPDUs within the A-MPDU causes the retransmission of only the destroyed part and not the entire A-MPDU. In addition, the greatest A-MPDU size changes adaptively on the basis of the FER and PER conditions of the link. The findings showed that while the FER can be kept within the allowable degree of VoIP services, the proposed technique performs better than the comparable technique in terms of throughput (higher throughput rate) and delay (less delay).

In summary, the VoIP aggregation technique over A-MPDU 802.11n has been discussed in this section. The proposed technique of Ref. [36] only uses the size threshold to control the A-MPDU size. The original 802.11n standard uses the size and delay thresholds. Thus, the proposed technique is unsuitable for delay-sensitive VoIP applications. In Refs. [32,39], the proposed technique resends only the damaged MSDUs, and thus, the shortcoming in Ref. [36] has been resolved. However, the proposed technique of Refs. [32,39] use the UD metric to control the A-MPDU size, which has a higher throughput rate and lower delay and packet loss compared with the comparable techniques. Thus, the proposed technique may be unsuitable in ensuring the same good result under different scenarios or other various link conditions, particularly because it only utilizes a single indicator (UD) to regulate the A-MPDU length. In Ref. [40],

the proposed technique retransmits only the destroyed MSDUs. Therefore, this scheme can address the weaknesses presented by Ref. [36]. However, similar to the work of Refs. [32,39], the proposed technique in Ref. [40] only utilizes a single parameter (FER) to adjust the A-MPDU length. Therefore, the obtained results may be inappropriate for dissimilar environments with different channel circumstances. The proposed technique in Refs. [14,33,37,38] uses various metrics (average medium access delay, end-to-end delay, 150 ms of tolerable delay, buffering delay thresholds, and congestion) to control the A-MPDU size. Thus, the scheme is suitable for different scenarios with various link conditions. Therefore, the proposed technique in Refs. [14,33,37,38] may be regarded as a superior existing VoIP aggregation technique over 802.11n. Specifically, Ref. [14] may have presented the best approach because of its much deeper analysis and more comprehensive results under different scenarios compared with those in Refs. [33,37,38]. Tab. 3 summarizes the aforementioned techniques.

Table 3: Adaptive and non-adaptive aggregation techniques

Technique	A-MPDU size determinant	Adopts 802.11e ACs	Preamble optimization	Partial retransmission	Approach
Ref. [14]	Different delay flavors congestion	Yes	No	Yes	Adaptive
Ref. [32]	Urgency delay	No	No	Yes	Adaptive
Ref. [33]	Different delay flavors congestion	Yes	No	Yes	Adaptive
Ref. [36]	Size	Yes	No	Yes	Non-adaptive
Ref. [37]	Different delay flavors congestion	Yes	No	Yes	Adaptive
Ref. [38]	Different delay flavors congestion	Yes	No	Yes	Adaptive
Ref. [39]	Urgency delay	No	No	Yes	Adaptive
Ref. [40]	FER BER	No	No	Yes	Adaptive

4 Discussion

This section presents the influences of the frame aggregation techniques on BW exploitation. This section also discusses an investigation of the preamble overhead in the implementation of the aggregation techniques. Tab. 4 presents the notations utilized in the equations and their sizes.

4.1 Impact of Frame Aggregation on BW Exploitation

The primary goal of frame aggregation is to improve the BW exploitation. On the one hand, aggregating numerous VoIP frames into a single A-MPDU prompts a significant approach of BW exploitation. The more the aggregated frames, the more the BW is utilized. Eqs. (1) and (2) are utilized to determine the proportion of the BW exploitation without and with frame aggregation, respectively. Figs. 8 and 9 show the BW exploitation with various payload sizes without and with frame aggregation, respectively. The data shown in Fig. 9 (with frame aggregation) are based on 20 bytes of payload size. A comparison between Figs. 8 and 9 indicate that the BW exploitation can be highly improved with frame aggregation. On the other hand, the 766 μ s after each VoIP frame will follow the A-MPDU frame with many VoIP frames. Thus,

more VoIP frames can be sent given the same channel time. The aggregation techniques can greatly enhance the BW exploitation [12,14,25,41,42].

$$BW_u \text{ Ratio} = \frac{P_s}{F_s} = \frac{P_s}{P_s + F_{hs}} \quad (1)$$

$$ABW_u = \frac{\sum_{i=1}^n P_s(i)}{AF_s} = \frac{\sum_{i=1}^n P_s(i)}{\sum_{i=1}^n P_s(i) + AF_{hs}} \quad (2)$$

Table 4: Equation notations

Notation	Description	Size (Byte)
ABW_u	Aggregated frame BW exploitation	N/A
AF_{hs}	Aggregated layer 2 frame preamble size	80 (theoretically)
AF_s	Aggregated layer 2 frame size	Varies
AP_{fs}	A-MPDU frame size	Varies
AP_{ho}	A-MPDU preamble overhead	N/A
AP_{hs}	A-MPDU preamble size	Varies
BW_u	BW exploitation	N/A
FCS	Frame check sequence size	4
F_{hs}	Layer 2 frame preamble size	80
F_s	Layer 2 frame size	Varies
M_h	MAC preamble size	40
P_s	Payload size (codec frame size)	10, 20, and 30
SF_d	Sub-frame delimiter	4
SF_p	Sub-frame padding size	0..3
SF_{pl}	Sub-frame payload size	50, 60, 70, (P_s + RTP/UDP/IP size (40))

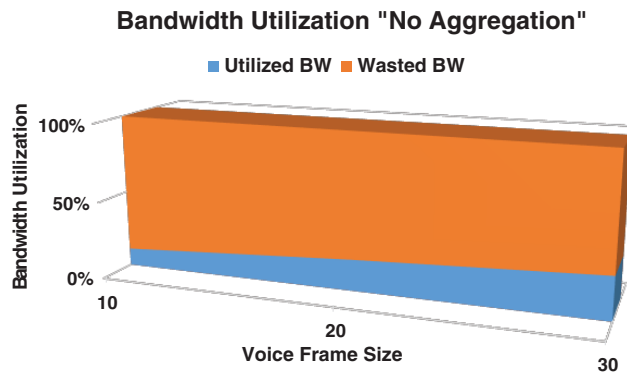


Figure 8: BW exploitation ratio without aggregation

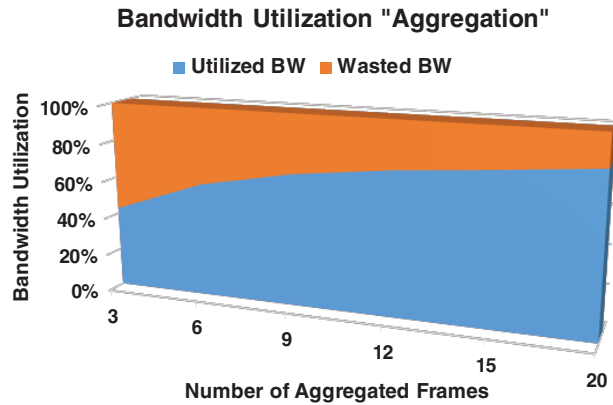


Figure 9: BW exploitation ratio with aggregation

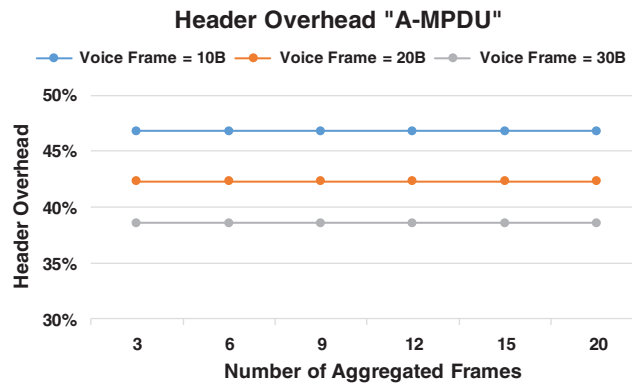


Figure 10: A-MSDU preamble overhead ratio

4.2 A-MPDU Preamble Overhead Analysis

The common payload of the VoIP packet is somewhat small. In this subsection, the preamble size of the A-MPDU is investigated by grouping dissimilar amounts and sizes of frames. The A-MPDU preamble overhead ratio is determined using Eq. (3).

$$AP_{ho} \text{ Ratio} = \frac{AP_{hs}}{AP_{fs}} = \frac{\sum_{i=1}^n SF_d(i) + SF_p(i) + M_h(i) + FCS(i)}{\sum_{i=1}^n SF_d(i) + SF_p(i) + M_h(i) + FCS(i) + SF_{pl}(i)} \quad (3)$$

In view of simplifying the calculation, the padding is set to be equal to 0, and the sub-frame size is always in multiples of 4. Fig. 10 shows the HOR with the A-MPDU aggregation technique. For voice frame sizes of 10, 20, and 30 bytes, the HOR begins from approximately 46.8%, 42.3%, and 38.6% when grouped into three sub-frames, respectively. The results indicate that the number of aggregated sub-frames does not affect the HOR, as HOR is only affected by the size of the voice frame. This finding can be attributed to the header that is added with each frame. However, the overhead resulting from the PHY layer timing interval is reduced when the A-MPDU aggregation technique is used [31].

5 Recommendations for Designing Aggregation Techniques and Possible Research Tips

An effective aggregation technique should follow a set of criteria for maximizing the BW exploitation ratio while maintaining an acceptable call quality. One of the key criteria is to select between adaptive and non-adaptive categories. As opposed to the non-adaptive techniques, the adaptive techniques tune the

A-MPDU size that is dependent on the current channel status. Therefore, the adaptive aggregation techniques generate a suitable A-MPDU size, consequently achieving better BW exploitation, compared with the non-adaptive aggregation techniques. Another key criterion is the selection of suitable parameters for tuning the A-MPDU size that is dependent on the current channel status in the adaptive aggregation techniques. Selecting the suitable parameters lessens the number of retransmissions, minimizes the delays, and maximizes the throughputs [1,12,14,33,34]. An effective aggregation technique should follow the aforementioned criteria in the design. However, the other criteria mentioned by Ref. [16] are not followed in the current techniques. In addition, the following criteria may be considered in future research:

1. Artificial intelligence (AI) is a computer science domain that has been widely integrated into many other sciences. Fuzzy logic (FL), as a major field in AI, is adopted by these sciences to provide intelligent-controlled supervision. FL is specifically suitable in environments requiring a set of certain criteria, including frequently changing parameters, nonlinearity, inaccurate measures, and complex models. FL provides intelligent-controlled supervision based on expert knowledge that can be described as approaching the decisions made by humans. For computer networks in general, FL is vastly adopted (e.g., network congestion control [43,44]) and has been proven to provide controlled supervision with noticeable performance improvements and stabilized environments. The studies should investigate accordingly the usefulness of integrating FL with frame aggregation techniques over 802.11n networks.
2. The existing aggregation techniques utilize numerous parameters for channel condition assessment, including congestion, delay, BER, FER, and PER. However, other channel condition assessment indicators can be utilized with the adaptive aggregation technique. Examples include SNR, jitter, packet loss, SNR, and interference ratio. These indicators should be examined and deliberated when designing an equation for the adaptive aggregation technique to be able to assess the channel conditions and calculate the suitable A-MPDU size [1].

6 Conclusion

This study has investigated the VoIP aggregation techniques over A-MPDU 802.11n networks in terms of BW exploitation and VoIP performance. The findings from our investigation indicate that the current VoIP aggregation techniques can greatly improve the BW exploitation. However, although some of the current aggregation techniques have considered the VoIP performance issue, much work is needed in this aspect. This study offers a set of primary guidelines that can be followed when designing a new aggregation technique. In addition, many other areas and techniques should be investigated with the frame aggregation techniques, including the implementation of AI techniques with frame aggregation techniques and the examination of new channel condition parameters, such as SNR, jitter, packet loss, and interference ratio. Several aggregation techniques can be designed in the future by using FL with different channel metrics. These techniques may be compared in view of selecting the most superior type, which then can be adopted for VoIP performance enhancement.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] M. M. Abualhaj, M. Kolhar, K. Qaddoum and A. A. Abu-Shareha, "Multiplexing voip packets over wireless mesh networks: A survey," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 8, pp. 3728–3752, 2016.
- [2] P. Christian <https://blog.telegeography.com/voice-traffics-slump-continued-in-a-big-way-last-year>, *Sept 2019*.

- [3] C. Hui Wang and Y. Shun Liu, "A dependable privacy protection for end-to-end voip via elliptic-curve diffie-hellman and dynamic key changes," *Journal of Network and Computer Applications*, vol. 34, no. 5, pp. 1545–1556, 2011.
- [4] R. Rajashekar, M. Di Renzo, K. Hari and L. Hanzo, "A beamforming-aided full-diversity scheme for low-altitude air-to-ground communication systems operating with limited feedback," *Journal of IEEE Transactions on Communications*, vol. 66, no.12, pp. 6602–6613, 2018.
- [5] Z. Farej and M. Jasim, "Investigation on the performance of the ieee802. 11n based wireless networks for multimedia services," in *Proc. of 2nd Int. Conf. for Engineering, Technology and Sciences of Al-Kitab*, Baghdad, Iraq, pp. 48–53, 2018.
- [6] D. Hucaby, *In Ccna Wireless 200–355 Official Cert Guide*, Cisco Press, Indiana, United States, 2016.
- [7] P. Dhawankar, H. Le-Minh and N. Aslam, "Throughput and range performance investigation for ieee 802.11 a, 802.11 n and 802.11 ac technologies in an on-campus heterogeneous network environment," in *Proc. of 11th Int. Symposium on Communication Systems, Networks & Digital Signal Processing*, Budapest, Hungar, pp. 1–6, 2018.
- [8] S. Seytnazarov, J. Ghoo Choi and Y. Tak Kim, "Enhanced mathematical modeling of aggregation-enabled wlans with compressed blockack," *Journal of IEEE Transactions on Mobile Computing*, vol. 18, no. 6, pp. 1260–1273, 2018.
- [9] K. Hassine and M. Frikha, "A voip focused frame aggregation in wireless local area networks: features and performance characteristics," in *Proc. of 13th Int. Wireless Communications and Mobile Computing Conf.*, Valencia, Spain, pp. 1375–1382, 2017.
- [10] C. Olariu, J. Fitzpatrick, Y. Ghamri-Doudane and L. Murphy, "A delay-aware packet prioritisation mechanism for voice over ip in wireless mesh networks," in *Proc. of IEEE Wireless Communications and Networking Conf.*, Doha, Qatar, pp. 1–7, 2016.
- [11] H. Natarajan, S. Diggi, M. R. Kanagarathinam, S. Kumar Srivastava and C. Bharti, "D-vowifi-a guaranteed bit rate scheduling for vowifi in non-dedicated channel," in *Proc. of 16th IEEE Annual Consumer Communications & Networking Conf.*, Las Vegas, NV, USA, pp. 1–6, 2019.
- [12] C. Vulkan, A. Rakos, Z. Vincze and A. Drozdy, "Reducing overhead on voice traffic," U.S. Patent No. 8,824,304, 2 Sep, 2014.
- [13] P. Fortuna and M. Ricardo, "Header compressed voip in ieee 802.11," *Journal of IEEE Wireless Communications*, vol. 16, no. 3, pp. 69–75, 2009.
- [14] S. Seytnazarov and K. Young-Tak, "Qos-aware adaptive a-mpdu aggregation scheduler for voice traffic in aggregation-enabled high throughput wlans," *Journal of IEEE Transactions on Mobile Computing*, vol. 16, no.10, pp. 2862–2875, 2017.
- [15] K. Sandlund, G. Pelletier and L. E. Jonsson, "The robust header compression (rohc) framework," No. RFC 5795, 2010.
- [16] Abualhaj, M. M., Hussein A. H., Kolhar M., and Abu AlHija M., "Survey and analysis of VoIP frame aggregation methods over A-MSDU IEEE 802.11 n wireless networks." *Computers, Materials & Continua*, vol. 66, no. 4, pp. 1283–1300, 2020.
- [17] J. Holub, M. Wallbaum, N. Smith and H. Avetisyan, "Analysis of the dependency of call duration on the quality of voip calls," *Journal of IEEE Wireless Communications Letters*, vol. 7, no. 4, pp. 638–641, 2018.
- [18] N. Gupta, K. Naresh and K. Harish, "Comparative analysis of voice codecs over different environment scenarios in voip," in *Proc. of Second Int. Conf. on Intelligent Computing and Control Systems*, Madurai, India, pp. 540–544, 2018.
- [19] M. O. Ortega, G. C. Altamirano, C. L. Barros and M. F. Abad, "Comparison between the real and theoretical values of the technical parameters of the voip codecs," in *Proc. of IEEE Colombian Conf. on Communications and Computing*, Barranquilla, Colombia, pp. 1–6, 2019.

- [20] Y. Niu, C. Wu, L. Wei, B. Liu and J. Cai, "Backfill: an efficient header compression scheme for open flow network with satellite links," in *Proc. of Int. Conf. on Networking and Network Applications*, Hakodate, Japan, pp. 202–205, 2016.
- [21] M. Abualhaj, "Ittp-mux: An efficient Multiplexing mechanism to improve voip applications bandwidth utilization," *International Journal of Innovative Computing, Information and Control*, vol. 15, no. 3, pp. 2063–2073, 2015.
- [22] IEEE Computer Society, "Part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications," 2012.
- [23] S. Yun, H. Kim, H. Lee and I. Kang, "100+ voip calls on 802.11 b: The power of combining voice frame aggregation and uplink-downlink bandwidth control in wireless lans.," *Journal of IEEE Journal on Selected Areas in Communications*, vol. 25, no. 4, pp. 689–698, 2007.
- [24] M. S. Gast, "802.11n a survival guide," *O'Reilly Media*, Sebastopol, CA, USA, 2012.
- [25] E. Charfi, C. Lamia and K. Lotfi, "Phy/mac enhancements and qos mechanisms for very high throughput wlans: A survey," *Journal of IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1714–1735, 2013.
- [26] M. Abu-Alhaj, M. Kolhar, L. Chandra, O. Abouabdalla and A. Manasrah, "Delta-Multiplexing: a novel technique to improve voip bandwidth utilization between voip gateways," in *Proc. of 10th IEEE Int. Conf. on Computer and Information Technology*, Bradford, UK, pp. 329–335, 2010.
- [27] D. Jianhua, "An adaptive packet aggregation algorithm (aam) for wireless networks," *PhD dissertation*, 2013.
- [28] K. Hassine, "Performance study of future wireless networks ieee 802.11 xy", *PhD dissertation*, 2016.
- [29] R. Karmakar, C. Samiran and C. Sandip, "Impact of ieee 802.11 n/ac phy/mac high throughput enhancements on transport and application protocols-a survey," *Journal of IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2050–2091, 2017.
- [30] A. Saif, M. Othman, S. Subramaniam, H. Abdul and A. Nor, "Frame aggregation in wireless networks: Techniques and issues," *Journal of IETE Institution of Electronics and Telecommunication Engineers Technical Review*, vol. 28, no. 4, pp. 336–350, 2011.
- [31] E. Charfi, C. Gueguen, L. Chaari, B. Cousin and L. Kamoun, "Dynamic frame aggregation scheduler for multimedia applications in ieee 802.11 n networks," *Journal of Transactions on Emerging Telecommunications Technologies*, vol. 28, no. 2, pp. e2942, 2017.
- [32] S. Seytnazarov and Y. T. Kim, "Qos-aware adaptive a-mpdu aggregation scheduler for enhanced voip capacity over aggregation-enabled wlans," *IEEE/IFIP Network Operations and Management Symposium*, Taipei, Taiwan, pp. 1–7, 2018.
- [33] Z. D. Dudu, "Packet aggregation for voice over internet protocol on wireless mesh networks," *PhD dissertation*, 2012.
- [34] J. M. Okech, M. O. Odhiambo and A. Kurien, "Packet voip aggregation: A mechanism to improve the performance of voip in wireless mesh networks (wmns)," *IST Transactions of Electrical and Electronic Systems-Theory and Applications*, vol. 1, no. 2, pp. 28–36, 2012.
- [35] J. M. Okech, M. O. Odhiambo and A. Kurien, "Packet voip aggregation: A mechanism to improve the performance of voip in wireless mesh networks (wmns)," *IST Transactions of Electrical and Electronic Systems-Theory and Applications*, vol. 1, no. 2, pp. 28–36, 2012.
- [36] M. G. Sarret, J. S. Ashta, P. Mogensen, D. Catania and A. F. Cattoni, "A multi-qos aggregation mechanism for improved fairness in wlan," in *Proc. of IEEE 78th Vehicular Technology Conference (Vtc Fall)*, Las Vegas, NV, USA, pp. 1–5, 2013.
- [37] S. Seytnazarov and Y. T. Kim, "Qos-aware adaptive mpdu aggregation of voip traffic on ieee 802.11 n wlans," in *Proc. of Int. Conf. on Network and Service Management (CNSM) and Workshop*, Rio de Janeiro, Brazil, pp. 356–359, 2014.
- [38] S. Seytnazarov and Y. T. Kim, "Qos-aware mpdu aggregation of ieee 802.11 n wlans for voip services," in *Proc. of the 2014 Int. Conf. on Electronics and Communication Systems II (ECS'14)*, Prague, Czech Republic, pp. 64–71, 2014.

- [39] E. Charfi, L. Chaari and L. Kamoun, "Joint urgency delay scheduler and adaptive aggregation technique in ieee 802.11 n networks," in *Proc. of 5th Int. Conf. on Communications and Networking (COMNET)*, Tunis, Tunisia, pp. 1–6, 2015.
- [40] M. Moh, M. Teng-Sheng and K. Chan, "Error-sensitive adaptive frame aggregation in 802.11 n wlan," in *Proc. of Int. Conf. on Wired/Wireless Internet Communications*, Berlin, German, pp. 64–76, 2010.
- [41] F. Blanco, D. Wing, J. Navajas, M. Perumal and J. Saldana, "Tunneling compressing and aggregation (tcm) traffic flows. reference model," 2015.
- [42] X. Jun and L. Feng, "Analysis and simulation for voip capacity in ieee 802.11 wlan," *Journal of Computational Information Systems*, vol. 8, no. 19, pp. 7955–7962, 2012.
- [43] M. M. Abualhaj, A. A. Abu-Shareha and M. M. Al-Tahrawi. "FLRED: An efficient fuzzy logic based network congestion control method," *Journal of Neural Computing and Applications*, vol. 30, no. 3, pp. 925–935, 2018.
- [44] M. Baklizi, H. Abdel-Jaber, A. Abu-Shareha, M. M. Abualhaj and S. Ramadass, "Fuzzy logic controller of gentle random early detection based on average queue length and delay rate," *International Journal of Fuzzy Systems*, vol. 16, no.1, pp. 9–19, 2014.