

Survey and Analysis of VoIP Frame Aggregation Methods over A-MSDU IEEE 802.11n Wireless Networks

Mosleh M. Abualhaj^{1,*}, Abdelrahman H. Hussein¹, Manjur Kolhar² and Mwaffaq Abu AlHija¹

¹Al-Ahliyya Amman University, Amman, Jordan

²Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia

*Corresponding Author: Mosleh M. Abualhaj. Email: m.abualhaj@ammanu.edu.jo

Received: 21 July 2020; Accepted: 11 September 2020

Abstract: The IEEE 802.11n standard has provided prominent features that greatly contribute to ubiquitous wireless networks. Over the last ten years, voice over IP (VoIP) has become widespread around the globe owing to its low-cost or even free call rate. The combination of these technologies (VoIP and wireless) has become desirable and inevitable for organizations. However, VoIP faces a bandwidth utilization issue when working with 802.11 wireless networks. The bandwidth utilization is inefficient on the grounds that (i) 80 bytes of 802.11/RTP/UDP/IP header is appended to 10–730 bytes of VoIP payload and (ii) 765 μ s waiting intervals follow each 802.11 VoIP frame. Without considering the quality requirements of a VoIP call, including frame aggregation in the IEEE 802.11n standard has been suggested as a solution for the bandwidth utilization issue. Consequently, several aggregation methods have been proposed to handle the quality requirements of VoIP calls when carried over an IEEE 802.11n wireless network. In this survey, we analyze the existing aggregation methods of VoIP over the A-MSDU IEEE 802.11n wireless standard. The survey provides researchers with a detailed analysis of the bandwidth utilization issue concerning the A-MSDU 802.11n standard, discussion of the main approaches of frame aggregation methods and existing aggregation methods, elaboration of the impact of frame aggregation methods on network performance and VoIP call quality, and suggestion of new areas to be investigated in conjunction with frame aggregation. The survey contributes by offering guidelines to design an appropriate, reliable, and robust aggregation method of VoIP over 802.11n standard.

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

1 Introduction

Over the past ten years, voice over IP (VoIP) has been increasingly ubiquitous with the proliferation of VoIP applications, such as Skype, Facebook Messenger, Viber, and so on. The Internet transferred more than 156 petabytes of VoIP data every month in 2015 [1]. In each month in 2018, the active users on Facebook



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.

Messenger reached 1.3 billion [2]. This tremendous ubiquity of VoIP has many incentives, the most important being the free or low-cost call rate [1,3].

In a different context, the wireless local area network (WLAN), particularly IEEE 802.11 standard, has spread everywhere, including schools, universities, shopping malls, and companies. The main driver of the wide spread is the ease of installation, especially in difficult-to-wire areas [4,5]. Several amendments have been made on IEEE 802.11 standard since its emergence. IEEE 802.11n, which has been standardized in September 2009, is a milestone among the IEEE 802.11 standard amendments [5,6]. In particular, the medium access control (MAC) layers of the previous 802.11 standards has undergone several enhancements. The main improvements in MAC layer are block acknowledgement (BA) and frame aggregation. BA acknowledges several data frames using only one acknowledge (ACK) frame, and up to 64 frames can be ACKed with one BA. A separate ACK frame is not required for each data frame. Frame aggregation, which is the main concept in this study, combines multiple data frames in one large frame [4,7,8]. The frame aggregation technique is discussed in detail in the subsequent sections. The propagation of VoIP and 802.11, along with the enhancements proposed in 802.11n, attracts all sectors to deploy VoIP over 802.11 wireless standard [5,9].

However, two key problems emerge with the aggregation of VoIP over 802.11 wireless networks. First, the voice quality of service (QoS) is degraded because of the increase in packet loss, delay, and jitter. The metrics (packet loss, delay, and jitter) of the 802.11 wireless networks are increased because of the scarcity of bandwidth, exposed links to a considerable amount of interference, increase in the contention of the VoIP packet and transmission-control protocol (TCP) traffic due to the growth of TCP applications, and increase in the access delay (the duration from the moment the packet reaches the access point until it is sent over the 802.11 wireless network) [1,10,11]. Apart from QoS, a vital problem is the extremely high overhead that leads to wasted bandwidth and resources, thereby reducing the 802.11 wireless synchronous call capacity (as discussed in the following section) [12–14]. This study focuses on the large header overhead with less concentration on the QoS problem.

The problem of large VoIP packet header overhead has been addressed by various methods. One of which is VoIP packet aggregation. Packet aggregation methods assemble many packets in one large packet with a single header, so each packet does not require a separate header [12,15,16]. The two aggregation methods proposed as part of the 802.11n standard are MAC service data unit (A-MSDU) and aggregation MAC protocol data unit (A-MPDU). This study aims to survey the current A-MSDU packet aggregation methods of VoIP packets over the IEEE 802.11n standard. The study discusses and emphasizes the inefficient bandwidth utilization of VoIP over A-MSDU 802.11n. It also demonstrates the effectiveness of current packet A-MSDU aggregation methods in enhancing the bandwidth utilization of 802.11n and their impact on VoIP QoS. Moreover, the study proposes general guidelines for building a robust aggregation method over 802.11n. This work, to the best of our knowledge, is the first to discuss aggregation methods of VoIP over A-MSDU 802.11n.

The organization of this article is as follows. Section 1 is the introduction wherein the objectives of the study are specified. Section 2 covers the main topics that explain the combination of VoIP and IEEE 802.11n wireless networks. Section 3 discusses the main aggregation approaches. Section 4 discusses the existing A-MSDU VoIP aggregation methods over IEEE 802.11n wireless networks. Section 5 discusses the impact of frame aggregation on bandwidth utilization and VoIP quality, analyzes the header overhead of aggregation A-MSDU, and elaborates the frame aggregation obstacles. Section 6 equips the researchers with the necessary guidelines to design a robust aggregation method and offers possible future research directions. Section 7 concludes the study.

2 Basic Concepts

To ensure that readers are well versed with the topic, this section clarifies the basic concepts connected to this work. These are VoIP over IEEE802.11n and the VoIP packet aggregation methods.

2.1 VoIP over 802.11n

VoIP technology increasingly dominates the telecommunications world by allowing the voice call to travel over an IP network. The quality of the VoIP is highly affected by packet loss and delay in that the longer the delay and the more the packet loss, the poorer the voice call quality [17]. The highest share of the delay is produced when generating the voice frame (VoIP packet data). The voice frame generated by hardware or software is known as codec. The codec collects the analogue voice signal. The collected analogue signals are then converted to digital data. Subsequently, the resulting digital data are compacted to produce the voice frame. The frame must be generated within a codec frame period, where each codec has a specific period to generate a voice frame. The larger the codec period, the bigger the frame size, and the higher the delay. Thus, the voice call quality is poor. The smaller the codec period, the smaller the frame size, and the lesser the delay; thus, the voice call quality is satisfactory. Tab. 1 lists a few voice codecs [18,19]. Tab. 1 shows that the typical frame size is up to 30 bytes. For each generated voice frame, a 40-byte RTP/UDP/IP header (8-byte user datagram protocol, 12-byte real-time transport protocol, and 20-byte Internet protocol) is attached to constitute the VoIP packet. Another header is attached to the packet at the data link layer (DLL). The size of the DLL header varies based on the used protocol. In the case of 802.11n, the DLL adds another 40-byte header apart from the 40-byte RTP/UDP/IP header, so the total header size is 80 bytes [6,20,21]. Fig. 1 presents the 802.11n frame format.

Table 1: Some of the well-known VoIP codecs

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

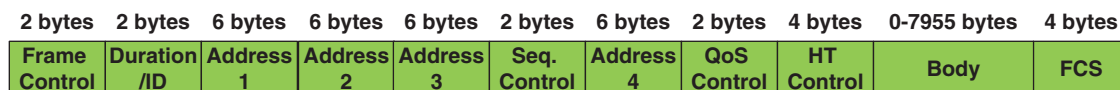


Figure 1: 802.11n frame format

On the one hand, the 802.11n 600Mb/s maximum speed is considered, and the time required to transmit an 80-byte VoIP packet header is $80 \times 8/600 = 1.06 \mu\text{s}$. In addition, the channel should be free prior to transmitting the VoIP packet. Therefore, prior to transmission, the VoIP packet should reach $765 \mu\text{s}$ under the following conditions: (i) The VoIP packet should reach $360 \mu\text{s}$, that is, $310 \mu\text{s}$ for backoff (BO) time and $50 \mu\text{s}$ for distributed inter-frame space (DIFS). (ii) Following the VoIP packet is $21 \mu\text{s}$, that is, $10 \mu\text{s}$ for short inter-frame space (SIFS) to ensure that the received packet is processed and $11 \mu\text{s}$ to ACK for the transmitted packet. (iii) At the PHY layer, $192 \mu\text{s}$ is added to each frame, that is, $144 \mu\text{s}$ at preamble and $48 \mu\text{s}$ at PHY layer convergence protocol (PLCP), and (iv) $192 \mu\text{s}$ is added to each ACK. The resulting overhead of these intervals ($765 \mu\text{s}$) and VoIP packet header ($1.06 \mu\text{s}$) is equal to approximately $766 \mu\text{s}$. On the other hand, the time required to transmit a 20-byte G.726 voice frame is $20 \times 8/600 = 0.27 \mu\text{s}$. The share of the voice frame from the bandwidth, which is calculated as the relative ratio between overhead time ($766 \mu\text{s}$) and voice frame time ($0.27 \mu\text{s}$), is negligible (approximately 0.000261)

[1,22,23]. The ratio varies based on the voice frame size. This approach causes great wastage of the available bandwidth. Fig. 2 shows the wasted bandwidth while sending a VoIP DLL frame given a 20-byte G.729.

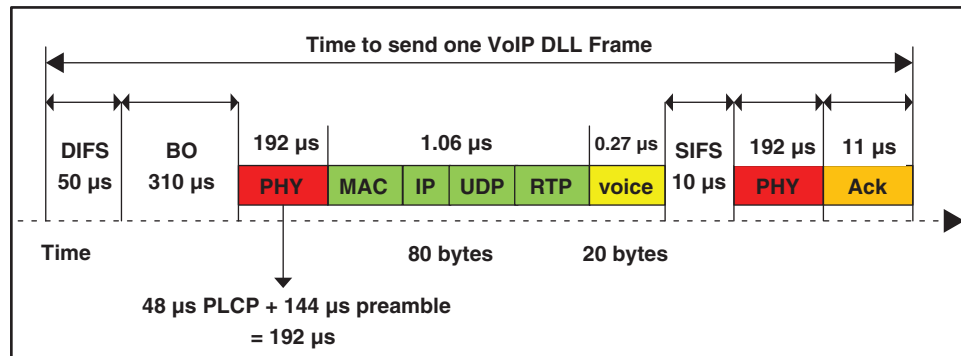


Figure 2: Time to send one VoIP DLL frame

As a result, one of the key improvements of the IEEE 802.11n standard is A-MSDU frame aggregation at DLL. Similar to packet aggregation, which is used to improve bandwidth utilization, DLL frame aggregation assembles many DLL frames in a large A-MSDU with only one header. The A-MSDU is then sent as a single burst through the PHY layer. Consequently, the ratio of the DLL frame payload to the header overhead is reduced, and more data are sent in a given transmit opportunity (TXOP). The airtime consumption is reduced because the channel contention overhead is reduced. In addition, the waiting timing intervals (DIFS, BO, and ACK) among many frames decrease because one large A-MSDU is sent instead of many DLL frames with a single channel access. Therefore, the bandwidth usage efficiency is highly enhanced and, thus, the capacity of the synchronous calls is increased [24–26].

Frame aggregation, in general, is one of the key methods used to reduce the large frame header overhead issue caused by carrying a VoIP packet over 802.11n. Fig. 3 illustrates the aggregation of three frames into one frame. The upper section of Fig. 3 exhibits the time required to send three frames separately. To send the three frames, a total of 2298 μs overhead (766×3) is needed. The lower section of Fig. 3 exhibits the time required to send the aggregated frame generated from the three frames. The aggregated frame only necessitates a virtual 766 μs. Thus, less overhead can be gained with the frame aggregation method. Aggregating more frames in one aggregated frame (large aggregated frame size is directly proportional to reducing the timing overhead and enhancing the bandwidth utilization [12,27].

2.2 A-MSDU IEEE 802.11n Aggregation Method

MSDU is the payload (layers 3 to 7) of the 802.11 data frame, whereas MPDU is the technical name of wireless frames and comprises a trailer, body (MSDU), and header. As previously mentioned, A-MSDU was put forward as aggregation method to be part of the IEEE 802.11n. The A-MSDU works at the top part of the MAC layer. A-MSDU allows a group of MSDUs transmitted to the same receiver to be encapsulated in one A-MSDU with a single 802.11n MAC header. Then, A-MSDU is passed down to the MAC sublayer to construct the MPDU, and then to the PHY layer to be sent to the channel. The aggregation process ends when the A-MSDU reaches the maximum size, based on the station capability, or when the first packet in the aggregation buffer reaches its delay limit. Therefore, the A-MSDU size should be selected and assigned carefully based on the link status, thereby causing difficulty and inspiring more studies. Each sub-frame, in the A-MSDU, contains three fields, namely, header (14 B), MSDU (0–2304 B), and padding (0–3 B) fields. The padding maintains the size of each MSDU at a multiple of four bytes to align de-aggregation at the side of the receiver. The sub-frame header is composed of the destination address

(6 B), source address (6 B), and length (2 B) fields. The header of the MAC layer is attached to the A-MSDU and sent as a single MPDU. Fig. 4 exhibits the A-MSDU frame structure. All sub-frames inside an A-MSDU must have the same transmitter and receiver addresses but not necessarily the same source and destination addresses. In addition, all sub-frames inside the A-MSDU must have the same QoS category, where voice, video, and best-effort traffic cannot be aggregated. The header size of the A-MSDU is relatively small compared to the traditional 802.11n without aggregation, thereby improving the header overhead, especially in the noiseless channel. One considerable issue in the A-MSDU method is that if any of the aggregated MSDUs is damaged, then the entire A-MSDU must be retransmitted. Therefore, the performance of A-MSDU is highly degraded especially in the erroneous environments [22,28–31].

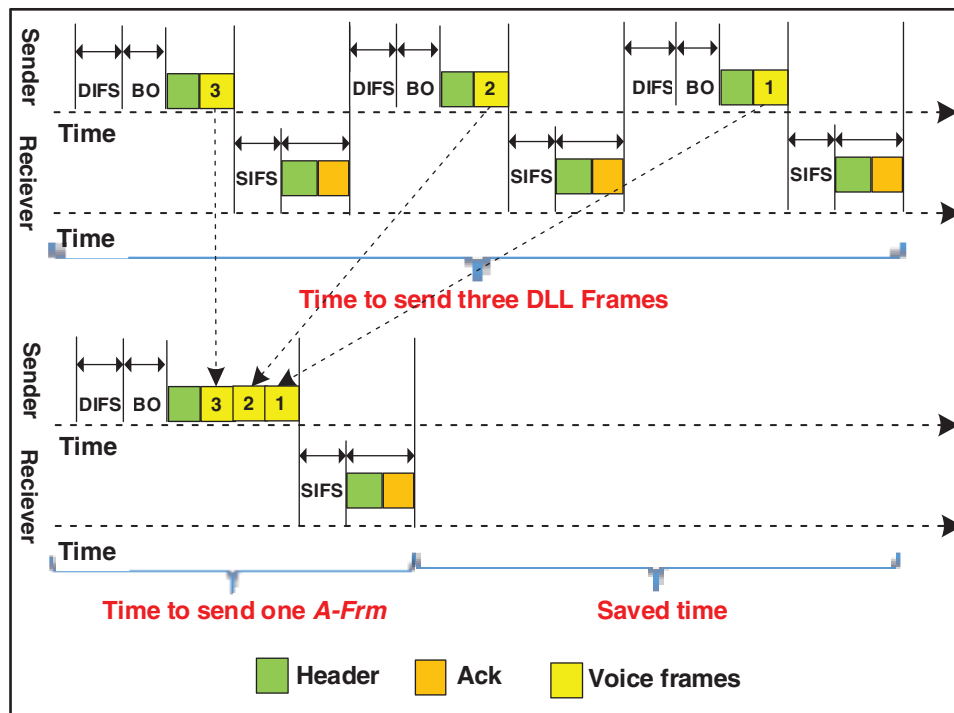


Figure 3: Reducing time intervals and header overhead by frame aggregation

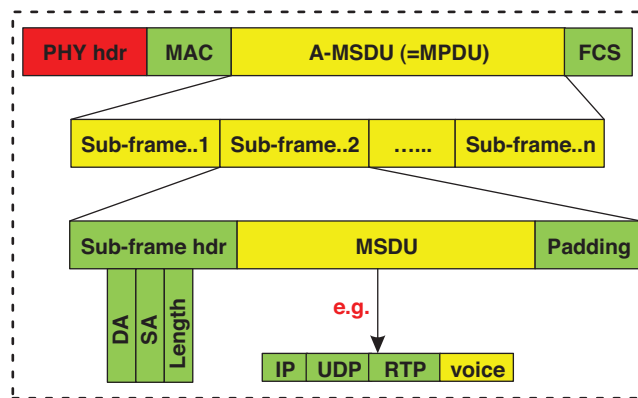


Figure 4: A-MSDU frame structure

For VoIP, although 802.11n provided higher throughput, the A-MSDU in the 802.11n standard has not considered the QoS requirements (packet loss, jitter, and maximum delay) of real-time applications including VoIP applications. As a result, the performance of 802.11n networks is severely degraded when the normal traffic is intermingled with the real-time VoIP traffic. This finding is caused by the MAC and PHY overhead linked with the small-size VoIP packet, the considerable decrease in the available throughput for the less priority traffics, and more loss of the VoIP packet due to the increase in the contention [14,32,33]. Therefore, several aggregation methods have been proposed to handle the conjunction between the A-MSDU and VoIP application traffic. The next section discusses these aggregation methods.

3 Aggregation Approaches

Various packet aggregation methods are available for VoIP over the IEEE 802.11n standard. The two common approaches are packet-level or frame-level methods. Packet-level aggregation methods are used for layer three and above, and the frame-level aggregation methods are used for layer two and below. This study focuses on the frame-level methods. Frame-level aggregation methods can be further categorized into several approaches, namely, A-MSDU, A-MPDU, adaptive, and non-adaptive methods. This section discusses the general idea of adaptive and non-adaptive aggregation approaches [31].

3.1 Adaptive Aggregation Approach

The adaptive aggregation approach controls the aggregated frame size. The aggregated frame size changes adaptively based on the channel quality metrics, such as signal-to-noise ratio (SNR), load, maximum delay, bit error rate (BER), congestion, and frame-error rate (FER). In meager quality channel, a large aggregated frame size is highly exposed to noise, interference, distortion, or bit synchronization errors, which increase the BER and FER ratios. Thus, the aggregated frame size has high probability to be corrupted. On the contrary, in the excellent quality channel, a large aggregated frame size reduces the header overhead and enhances the bandwidth utilization. The adaptive aggregation approach adjusts the aggregated frame size based on channel quality and network condition [12,34,35].

3.2 Non-Adaptive Aggregation Approach

The non-adaptive aggregation approach controls the aggregated frame size based on specific limits, including time period, the number of packets, and size. The factual channel quality condition is not considered. Therefore, the aggregated frame size does not improve the channel quality. The performance of the aggregation method is degraded [12,34].

4 VoIP Aggregation Methods over A-MSDU 802.11n

This section discusses the existing VoIP aggregation methods over A-MSDU 802.11n. Only the methods proposed after 802.11n became a standard were discussed; thus, the method in Li et al. [36] was not discussed here because it was published on April 2009 [6]. In addition, only the methods that considered the VoIP application requirements were discussed. We divided the existing methods into non-adaptive and adaptive.

4.1 Non-Adaptive Aggregation Methods

A non-adaptive frame aggregation method that works at the top part of the MAC layer has been developed [37–39]. The frames that share the same path to one receiver are assembled in a large A-MSDU frame. These assembled MSDUs must have the same QoS level. The proposed methods rely on two configurable metrics to control the A-MSDU and the aggregation process, namely, size and delay. The process ends when one of the two conditions is achieved; if either the size of the queued frame in the buffer reaches a preconfigured A-MSDU size threshold or if the first frame in the queue reaches a

preconfigured delay threshold. In addition to frame aggregation, the proposed method has reduced the header size of the aggregated sub-frames inside the A-MSDU by optimizing the duplicated fields in the sub-frame header. However, a new aggregation header with new fields is added to the aggregated frame structure. In addition, contrary to the traditional A-MSDU, the new A-MSDU structure allows retransmitting only the damaged MSDUs instead of the entire A-MSDU. Therefore, the network performance is highly improved. The evaluation analysis of the method in Refs. [37–39] generally showed that the proposed method is efficient especially with small-size MSDUs, such as VoIP MSDUs. In Saif et al. [38], the proposed method was implemented and tested in comparison with the traditional A-MSDU and A-MPDU under various scenarios and using MSDU sizes ranging from 64–1500 bytes. The evaluation result proved that the proposed method outperformed A-MSDU and A-MPDU. The header overhead ratio has been reduced by up to 12% and 57% compared with A-MSDU and A-MPDU, respectively. Regarding throughput, the proposed method has increasingly outperformed A-MSDU and A-MPDU with different BERs.

In Refs. [40,41], another non-adaptive frame aggregation method was proposed. Similar to the method in Refs. [37–39], the method in Refs. [40,41] works at the top part of the MAC layer as well as assembles the frames that share the same path to one receiver in a large A-MSDU frame. In addition, the proposed method adopted the service differentiation in the IEEE 802.11e standard and grouped the traffic into four access categories (ACs) with a predefined A-MSDU frame size threshold. For VoIP AC, the delay should not exceed the allowable maximum delay. The aggregation process is ended and the A-MSDU frame is sent to its destination when the size of the queued frame in the buffer reaches a predefined size threshold or when the delay reaches the allowable maximum delay. Similar to the traditional A-MSDU, when any of the MSDUs inside A-MSDU is damaged, the entire A-MSDU must be retransmitted, thereby degrading the performance, especially in the erroneous networks. The proposed method was implemented to evaluate the delay and throughput improvement. The results showed that the delay has been reduced and the throughput has been increased in the tested scenarios in terms of VoIP, thereby improving the overall performance.

In Refs. [42–44], a non-adaptive frame aggregation method similar to that in Refs. [37–39] to a large extent was proposed. This method combines frame aggregation and packet header optimization. On the one hand, the frames with the same destination and QoS level are grouped at the top part of the MAC layer in one large A-MSDU frame. The process is completed if either the size of the A-MSDU or the delay of the first frame in the buffer reaches a specific threshold. On the other hand, the header of each sub-frame within the A-MSDU is optimized and a new common aggregation header is added after the MAC header. The new header scheme allows the retransmission of only the damaged MSDUs within the A-MSDU, instead of the entire A-MSDU. In Maqhat et al. [44], the method has been extended to separate the voice traffic in the different queues from that used for non-voice traffic. The main difference between Refs. [37–39] and Refs. [42–44] is that the sub-frame header's size is 4-bytes and 16-bytes, respectively. This extra 12-bytes in size consumes more bandwidth; however, it does not affect the core features of the two methods. Similar to Refs. [37–39], the evaluation analysis of the method in Refs. [42–44] showed that the proposed method is efficient especially with small-size MSDUs, such as VoIP MSDUs. In addition, the retransmission of only the damaged MSDUs has highly improved network performance. In Saif et al. [43], the proposed method was implemented and tested compared with the traditional A-MSDU and A-MPDU with different scenarios and different MSDU sizes from 64–1500 bytes and two different data rates (150 and 300 Mbps). The evaluation result proved that the proposed method outperformed the A-MSDU and A-MPDU in the tested scenarios. The throughput has improved from 9% to 58%, depending on the MSDU size, compared with A-MSDU. In addition, the proposed method has outperformed the A-MSDU in terms of delay (imposes less delay). For A-MPDU, the throughput has improved from 2% to 15%, depending on the MSDU size.

In summary, this subsection examines the non-adaptive aggregation methods. In Refs. [40,41], similar to the original 802.11n standard, the proposed method uses the aggregated frame size and delay thresholds to control the aggregated frame size. To satisfy the VoIP application QoS requirements, the delay threshold was bounded by the maximum tolerable delay of VoIP, contrary to the specific threshold in the original 802.11n standard. However, similar to the typical 802.11n A-MSDU method, if any of the aggregated MSDUs is damaged, then the entire A-MSDU must be retransmitted. Therefore, the performance is highly degraded in erroneous environments. In Refs. [37–39,42–44], the parameters (aggregated frame size and delay thresholds) and method are the same as that proposed in the original 802.11n standard to control the aggregated frame size and end the aggregation operation. In addition, they have reduced the header size of the sub-frames to save more bandwidth. The new header structure allows to retransmit only the damaged MSDUs. Therefore, it improves the performance (increases the throughput and reduces the delay). The main difference between the methods in Refs. [37–39] and Refs. [42–44] is that the header size of the sub-frames is 4 and 16 bytes, respectively. The extra 12 bytes in Refs. [42–44] consumes more bandwidth. Therefore, the method in Refs. [37–39] is the best among the non-adaptive multiplexing methods. However, the non-adaptive aggregation methods typically assign a constant value to limit the aggregated frame size, which severely degrades the effectiveness of the aggregation method. In other words, the 802.11 network channel performance parameters (generally, interference and load) fluctuate frequently. Thus, when the channel performance parameters are excellent (soft interference and soft load), the aggregated frame size may enlarge, resulting in better bandwidth utilization. On the contrary, when the channel performance parameters are ineffective (high interference and high load), the aggregated frame size may reduce, thereby reducing the (i) delay induced from the aggregation process time and (ii) the packet loss caused by the large-size aggregated frame. Consequently, the VoIP call quality is enhanced. Accordingly, the aggregated frame size in frame aggregation methods should change adaptively based on the 802.11 network channel performance parameters. Such quality characteristic is available in the adaptive aggregation method, which is explained in the next subsection [12,34,35].

4.2 Adaptive Aggregation Methods

Shin et al. [45] proposed an adaptive frame aggregation method that works at the top part of the MAC layer. The frames that share the same path to one receiver are assembled in a large A-MSDU frame. Similar to the traditional A-MSDU, the damage of any of the aggregated MSDUs causes the retransmission of the entire A-MSDU. The maximum size of the A-MSDU changes adaptively based on a new derived packet error rate (PER) algorithm, which estimates the quality of the link. Simulation scenarios focus on QoS-sensitive applications with small frames, such as VoIP. The simulation result showed that the proposed method is effective in terms of system performance improvement and that it is a desirable solution, especially under noisy channels.

Similar to Shin et al. [45], Yeon et al. [46] proposed an adaptive frame aggregation method that works at the top part of the MAC layer and assembles the frames that share the same path to the same receiver in a single large A-MSDU frame. Moreover, the damage of any of the aggregated MSDUs causes the retransmission of the entire A-MSDU. In addition, the proposed method adopted the service differentiation in the IEEE 802.11e standard and grouped the traffic into four ACs. The A-MSDU frame size changes adaptively based on monitoring the channel BER, calculating the optimal frame waiting time in the buffer, and determining the acceptable FER rate based on type application and the dynamic aggregation and fragmentation (DAF) algorithm. DAF fragments the packets received from upper layers. Compared with large ones, small frames are less influenced by the channel error. These components are used together in the proposed method to calculate the suitable A-MSDU frame size. Theoretical and simulation analysis of the proposed method was performed. The result showed that the delay was within

the acceptable range, and the throughput was increased in the tested scenarios. For instance, the throughput improved more than twice compared with the conventional method under BER of 10^{-4} .

Different from the aforementioned methods, the proposed method in Moh et al. [47] implements two-level frame aggregation. It combines several MSDUs intended to one receiver in one A-MSDU. Then, it combines several A-MSDUs in one A-MPDU. In the traditional A-MPDU, the damage of any of the aggregated MPDUs causes the retransmission of that MPDU alone. The maximum A-MPDU size changes adaptively based on a specific equation that calculates the tolerable FER with the current BER as basis. The findings indicated that the proposed method outperformed the comparable method in terms of delay (less delay) and throughput (higher throughput rate), while maintaining the FER within the acceptable level of VoIP applications.

In summary, this subsection examines the adaptive aggregation methods. In Shin et al. [45], the proposed method has used a derived PER algorithm to control the aggregated frame size. In Yeon et al. [46], the proposed method has used four different criteria (monitoring BER, calculate the optimal frame waiting time in the buffer, determine the acceptable FER rate based on type application, and DAF algorithm) to control the aggregated frame size. Combining these four criteria in one method has provided a noticeable throughput improvement. However, Shin et al. [45] and Yeon et al. [46] have a common considerable shortcoming, that is, if any of the aggregated MSDUs are damaged, then the entire A-MSDU must be retransmitted. Therefore, the performance is highly degraded in erroneous environments. In Moh et al. [47], the proposed method retransmits only the damaged MSDUs. Thus, it handles the shortcomings in Shin et al. [45] and Yeon et al. [46]. However, the proposed method in Moh et al. [47] only uses one metric (FER) to control the aggregated frame size. Thus, the acquired results may be unsuitable for different scenarios with various link conditions.

The present methods to aggregate the VoIP packet over A-MSDU 802.11n are introduced in this section. These methods are classified into non-adaptive and adaptive. Specific thresholds, such as delay and size, are used to determine the aggregated frame size in non-adaptive methods. In adaptive methods, the link quality parameters, such as delay with different flavors, size, PER, FER, BER, and congestion, are used to determine the aggregated frame size. Tab. 2 lists the discussed methods.

Table 2: List of adaptive and non-adaptive aggregation methods

Method	Approach	Aggregated frame size determinator	Header optimization	Partial retransmission	Adopts 802.11e ACs
Ref. [37]	Non-adaptive	Size Configured Delay	Yes	Yes	No
Ref. [38]	Non-adaptive	Size Configured Delay	Yes	Yes	No
Ref. [39]	Non-adaptive	Size Configured Delay	Yes	Yes	No
Ref. [40]	Non-adaptive	Size 150 ms Delay	No	No	Yes
Ref. [41]	Non-adaptive	Size 150 ms Delay	No	No	Yes
Ref. [42]	Non-adaptive	Size Configured Delay	Yes	Yes	No
Ref. [43]	Non-adaptive	Size Configured Delay	Yes	Yes	No
Ref. [44]	Non-adaptive	Size Configured Delay	Yes	Yes	No
Ref. [46]	Adaptive	PER	No	No	No
Ref. [47]	Adaptive	BER Buffer Delay FER DAF	No	No	Yes
Ref. [51]	Adaptive	FER BER	No	Yes	No

5 Discussion

This section discusses the impact of frame A-MSDU aggregation methods on bandwidth utilization and VoIP quality analyses of the header overhead when applying the aggregation methods. Moreover, the obstacle faced by the frame aggregation methods is demystified. [Tab. 3](#) lists the notations used in this section and their corresponding sizes.

Table 3: Equations' notations

Notation	Description	Size (Byte)
BW_u	Bandwidth Utilization	N/A
P_s	Payload size (codec frame size)	10, 20, and 30
F_s	DLL frame size	Varies
F_{hs}	DLL frame header size	80
ABW_u	Aggregated Frame Bandwidth Utilization	N/A
AF_s	Aggregated DLL frame size	Varies
AF_{hs}	Aggregated DLL frame header size	80 (theoretically)
AS_{ho}	A-MSDU header overhead	N/A
AS_{hs}	A-MSDU header size	Varies
AS_{fs}	A-MSDU frame size	Varies
M_h	MAC header size	40
FCS	Frame check sequence size	4
SF_h	Sub-frame header size	14
SF_{pl}	Sub-frame payload size	50,60,70, ($P_s + \text{RTP/UDP/IP size (40)}$)
SF_p	Sub-frame padding size	0.3
AS_s	A-MSDU size	Varies

5.1 Impact of Frame Aggregation on Bandwidth Utilization

Enhancing bandwidth utilization is the key objective of the frame aggregation methods of VoIP over 802.11n networks. Aggregation methods fulfill this objective on various aspects. For one, the traditional VoIP frame payload size is between 10 and 30 bytes, which is based on the used codec, but the header size attached to each frame reaches 80 bytes. [Eq. \(1\)](#) can be used to calculate the ratio of the bandwidth utilization. The bandwidth utilization is approximately between 11.1% and 27.3% of the available bandwidth based on the codec's frame size, thereby wasting a considerable amount of bandwidth. [Fig. 5](#) illustrates the bandwidth utilization with different payload sizes. Aggregating many VoIP frames together in one header leads to substantial bandwidth saving in accordance with the number of aggregated VoIP frames inside one aggregated frame. [Eq. \(2\)](#) can be used to calculate the theoretical ratio of bandwidth utilization with frame aggregation. [Fig. 6](#) presents the theoretical bandwidth utilization with a different number of aggregated DLL frames when the codec frame size is equal to 20 bytes. The A-MSDU sub-frame header is not considered here because it is discussed in the next subsection. For another, a delay of 766 μs follows each typical VoIP frame, as explained in Section 2.1. However, the same value (766 μs) follows many VoIP frames aggregated in one large aggregated frame, highly increasing the number of

transmitted frames on the same channel. Third, the channel capacity of ongoing calls at the same time increases due to the following reasons: (i) The saved bandwidth, resulting from aggregating multiple frames in one header tolerates the channel to carry more concurrent calls; (ii) Confining the 766 μ s to one large aggregated frame instead of several small frames reduces the channel occupancy time and tolerates the channel to carry more concurrent calls as well [12,14,25,48,49]. These aspects reflect the bandwidth utilization. Therefore, the aggregation methods greatly enhance the bandwidth utilization.

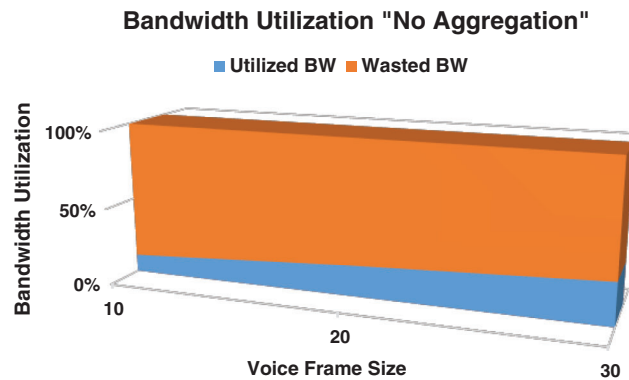


Figure 5: Bandwidth utilization ratio without aggregation

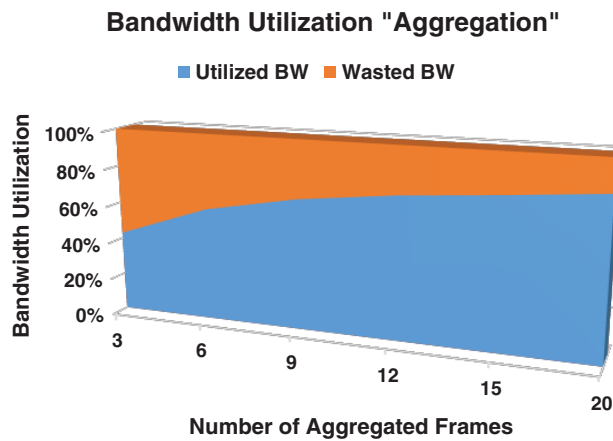


Figure 6: Bandwidth utilization ratio with aggregation

$$BW_u \text{ Ratio} = \frac{P_s}{F_s} = \frac{P_s}{P_s + F_{hs}} \tag{1}$$

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

5.2 A-MSDU Header Overhead Analysis

The typical VoIP packet payload is relatively small. Therefore, sufficient VoIP frames should exist in the buffer to obtain valuable aggregation through reducing the header size overhead, thereby resulting in delay. Here, we investigate the header size of A-MSDU by aggregating different numbers of frames with different sizes. The A-MSDU header overhead ratio can be calculated using Eq. (3).

$$\begin{aligned}
 AS_{ho} \text{ Ratio} &= \frac{AS_{hs}}{AS_{fs}} = \frac{M_h + FCS + \sum_{i=1}^n SF_h(i) + SF_p(i)}{M_s + FCS + AS_s} \\
 &= \frac{M_h + FCS + \sum_{i=1}^n SF_h(i) + SF_p(i)}{M_s + FCS + \sum_{i=1}^n SF_h(i) + SF_p(i) + SF_{pl}(i)}
 \end{aligned} \tag{3}$$

To simplify the calculation, we assume that the size of the sub-frame is always in multiples of 4; hence, the padding is equal zero. The header overhead ratio (HOR) was calculated with different voice frame sizes. Fig. 7 shows the HOR with the A-MSDU aggregation method. For 10, 20, and 30 bytes of voice frame sizes the HOR starts from approximately 35%, 31%, and 28%, respectively, when aggregating three sub-frames. The HOR decreases when increasing the number of aggregated sub-frames in a single A-MSDU.

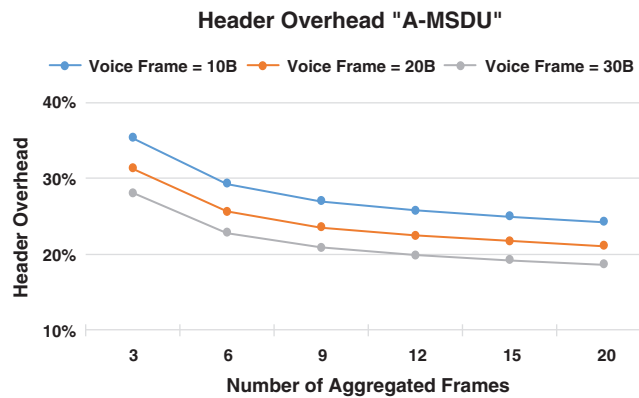


Figure 7: A-MSDU header overhead ratio

5.3 Impact of Frame Aggregation on VoIP Quality

The quality of the VoIP calls should satisfy certain requirements of the call parties. The three main parameters to measure VoIP call quality are packet loss, jitter, and delay. The delay of the VoIP call should not exceed 150 ms; otherwise, the quality of the call is degraded. The typical sources of delay for VoIP calls are packetization, transmission, processing, and queuing delay. The aggregation methods add two other sources, which are as follows: (i) The aggregation period while waiting for the VoIP frames to be gathered at the buffer and (ii) The time required to perform aggregation/de-aggregation. Delay is the main cause of using the frame aggregation methods to enhance bandwidth utilization. However, aggregation methods can recover the induced delay in two aspects. First, the queuing delay is decreased through handling and transmitting one large aggregated frame instead of several small frames. Second, the 766 μ s that follows each small frame is confined to one large aggregated frame instead. The aggregation methods can lengthen or shorten the delay. Thus, the proposed aggregation method should consider a mechanism that estimates the time to hold the frame in the buffer for aggregation and to indemnify the time [12,28,50–53]. The jitter should be low (not more than 30 ms) to obtain good call quality. The main sources of jitter are traveling the frames through various paths, waiting time at the queue, and contention. In case of aggregation, the jitter is slightly enhanced (reducing) due to several causes, as follows: (i) The contention is decreased because of the less number of frames in the channel, (ii) The frames must wait slightly to achieve aggregation, and (iii) One large aggregated frame travels through one path instead of several paths for several small frames [1,34,54]. Finally, for VoIP, packet loss occurs when the packets fail to reach the receiver or when the delay of the packets is more than the acceptable delay. The percentage of the tolerable lost frame is subjected to the frame size and codec. Various codecs use various packet loss concealment (PLC) mechanisms to replace a missing packet by

deriving a similar one. In general, up to 4% of packet loss is tolerable. In the case of aggregation methods, the packet loss is lessened due to the better utilization of the buffers. One header is added to the aggregated frame instead of several headers to the several small frames. Therefore, more frames may be saved in the buffer. Nevertheless, enlarging the aggregated frame size increases its possibility to be corrupted and lost, especially in noisy channels, where the packet loss percentage increases. Moreover, large aggregated frame is considered when several small frames corrupted, reducing the effectiveness of PLC mechanisms. Accordingly, selecting the suitable size of the aggregated frame is important [1,35,55]. Thus, a proposed aggregation method should consider a mechanism that estimates the number of frames to be aggregated into one aggregated frame. The proposed method in Gupta et al. [18] proved that aggregating VoIP DLL frames does not affect VoIP application QoS. Moreover, the aggregation method should use a well-designed mechanism that calculates the number of aggregated frames, the suitable aggregated frame size, and delay.

5.4 Obstacles of Frame Aggregation

Many issues hinder the aggregation of VoIP application packets when working with A-MSDU 802.11n networks. First, frame aggregation may debase the quality of VoIP conversion (increase in packet loss, jitter, and delay) if the aggregation method is not designed effectively, as discussed in Refs. [1,12,28,34,35,50–54]. Second, saving the bandwidth is preferred only if the links are saturated with the VoIP calls; otherwise, only a few number of frames is aggregated. Thus, saving bandwidth is either less favorable or unfavorable [1,38,53]. Third, aggregating several VoIP frames from different streams in one aggregated frame provides the same QoS level, which blocks the merit of traffic prioritization [1,53]. Finally, the processes of aggregation and de-aggregation burden the devices and, thus, require extra capability [1,55]. However, these hindrances may be alleviated by selecting convenient measures when designing an aggregating method. Designing a convenient aggregating method that enhances bandwidth utilization continues to challenge A-MSDU 802.11n developers. Tab. 4 lists the factors impacted by frame aggregation methods.

Table 4: Factors impacted by frame aggregation

Attribute	Influence
HOR	Grouping multiple frames in one header instead of a separate header to frame; thus, the HOR is reduced.
BWU	Reducing the HOR saves the BW, thereby improving the BWU ratio.
Capacity	A channel with the same available BW holds more concurrent calls due to (i) reduced HOR at the DLL and (ii) shortened timing interval at the PHY layer.
Delay	The delay can be increased depending on the time required to carry out aggregation/de-aggregation and the frame waiting time at the buffer. However, this delay can be compensated by reducing the queuing delay and shortening the 766 μ s delay to one frame.
Packet loss	Packet loss can be decreased because the buffer intakes more frames but large frames are more likely to be damaged and thus loss. Therefore, in error-free (or few errors) channels, packet loss decreases and in error-prone (or more errors) channels packet loss increases.
Jitter	Jitter is slightly enhanced due to the decrease in the frame waiting time at the buffer and because large aggregated frame travels through one path.

(Continued)

Table 4 (continued).

Attribute	Influence
Traffic	The amount of traffic running over a channel is decreased due to aggregate multiple frames in one aggregated frame.
Frame forwarding	Reducing the number of frames leads to improved frame forwarding process.
Network overload	Improving BWU and shortening the frame forwarding process leads to reduced network overload.
Congestion	Improving BWU, shortening the frame forwarding process, and reducing network overload leads to reduced congestion.
Hardware Capability	More hardware capability is required due to the aggregation/de-aggregation processes at the stations.

6 Guidelines for Designing Aggregation Methods and Potential Research Directions

The aggregation methods of VoIP packet over A-MSDU 802.11n aim to make the number of ongoing concurrent calls as large as possible to improve call quality or at least maintain the quality within the acceptable call requirements. To fulfill this purpose, a new aggregation method should be designed considering several issues. Among these issues is selecting what method to design (non-adaptive vs. adaptive) [34]. The aggregated frame size is the key criterion when designing an aggregation method. For non-adaptive aggregation methods, the aggregated frame size is confined by presetting the thresholds and does not consider the instantaneous link conditions. This approach may produce inefficient aggregated frame size and cause some problems as follows: (i) In unstable (erroneous) links, the large aggregated frame size increases packet loss, and (ii) In stable (little faults) links, the small aggregated frame size achieves inefficient bandwidth utilization. For adaptive aggregation methods, the aggregated frame size is adjusted adaptively by considering the instantaneous link conditions. Thus, the adaptive aggregation methods produce efficient aggregated frame size and, therefore, provides better bandwidth utilization and less packet loss than non-adaptive aggregation methods [12,14,33,34,55]. Thus, the adaptive aggregation methods are preferred over non-adaptive aggregation ones. Another important issue is that many link status evaluation parameters can be used to estimate the aggregated frame size in the adaptive aggregation method [1]. Selecting the appropriate parameters reduces the number of retransmission, maximize throughput, and minimize the delay. Therefore, an adaptive aggregation method should consider selecting the appropriate link status evaluation parameters and sending them to a robust equation and algorithm. Another vital issue is selecting the aggregating layer appropriately. Aggregation can occur at the various layers: Application, transport, network, DL, and PHY layer. Aggregation at the application layer achieves the highest bandwidth utilization, and the PHY layer achieves the lowest bandwidth utilization. This finding is due to the reduced header overhead of the aggregated frames [1,31,34,52]. A reliable and feasible aggregation method should consider these issues in the design. However, other aspects are not considered in the existing methods. The following aspects may cause the preference of an aggregation method over another, so they should be considered in future research.

1. Header compression is one of the key methods used to improve the exploitation of the bandwidth. The researchers have achieved a considerable compression of the IP/UDP/RTP header from 40 to 2 bytes. The VoIP frame aggregation methods combine several packets with separate IP/UDP/RTP header in one aggregated frame. Therefore, a new aggregation method that works with header compression can be developed. This method can compress the IP/UDP/RTP headers inside the

aggregated frame. Furthermore, the header compression can be extended to span the DLL, where the sub-frames' header inside the aggregated frame can be compressed in a similar approach [13,15].

2. The adaptive aggregation methods provide superior performance results to non-adaptive aggregation methods because they use channel status evaluation parameters to find the favorable aggregated frame size. The existing aggregation methods use several channel status evaluation parameters, including PER, FER, BER, delay with different flavors, and congestion. However, many other channel status evaluation parameters can be used with the adaptive aggregation method, such as packet loss, jitter, SNR, and signal-to-noise and interference ratio. These parameters should be investigated and considered when proposing a formula for the adaptive aggregation method to evaluate the channel status and find the favorable aggregated frame size [1].
3. QoS is the main factor when implementing VoIP technology. In 802.11n, a scheduling algorithm, as part of QoS, is not considered part of the standard and allowed to remain for the vendors. The existing proposed scheduling algorithms are restricted to IEEE 802.11e standard network; thus, they do not address the aggregation methods. However, implementing the scheduler in 802.11e with frame aggregation degrades the performance because of the small aggregation within the high priority ACs. In Abualhaj [16], an aggregation method that works in conjunction with a modified version of priority queuing (PQ) scheduler was proposed. However, appropriate scheduling algorithms that adapt with frame aggregation and satisfy the traffic quality requirements must be selected. More schedulers other than PQ should be implemented and tested with the aggregation method, including low latency queuing (LLQ), weighted fair queuing (WFQ), and class-based weighted fair queuing (CBWFQ) [31,32].

7 Conclusion

Several organizations have deployed VoIP over 802.11 wireless networks due to the features provided by these technologies. However, integrating these technologies faces a serious bandwidth utilization problem. 802.11n, which is a prominent amendment of 802.11 wireless, has proposed a frame aggregation method feature that aggregates several frames in one header. Nevertheless, the A-MSDU aggregation approach of 802.11n standard has not considered the QoS requirements (packet loss, jitter, maximum delay) of real-time VoIP applications. As a result, the researchers have proposed several VoIP frame aggregation methods that consider the VoIP call QoS requirements. This study has surveyed the existing VoIP aggregation methods over the A-MSDU IEEE 802.11n wireless standard. The sources that cause the bandwidth utilization problem when aggregating the VoIP frames over 802.11n standard are analyzed. The main approaches of frame aggregation methods and existing aggregation methods, the impact of the frame aggregation on VoIP application quality and bandwidth utilization, the guidelines to design a new robust aggregation method, and new study area and aspects that can be integrated with the frame aggregation to enhance quality and bandwidth utilization are discussed. In the future, the VoIP aggregation methods over other 802.11 amendments, such as 802.11ac and 802.11ad, will be investigated.

Acknowledgement: We appreciate the Research Deanship of Al-Ahliyya Amman University (AAU), Jordan and the Research Deanship of Prince Sattam Bin Abdulaziz University, Kingdom of Saudi Arabia, for providing research resources and equipment and various research programs to encourage research among faculty members.

Funding Statement: The author(s) 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, 2019. [Online]. Available: <https://blog.telegeography.com/voice-traffics-slump-continued-in-a-big-way-last-year>
- [3] C. H. Wang and Y. S. 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, *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 802.11a, 802.11n and 802.11ac technologies in an on-campus heterogeneous network environment," in *Proc. of 11th Int. Sym. on Communication Systems, Networks & Digital Signal Processing*, Budapest, Hungary, pp. 1–6, 2018.
- [8] S. Seytnazarov, J. G. Choi and Y. T. 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, J. 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.*, 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, 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 Y. T. Kim, "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] M. Abualhaj, "CA-ITTP: An efficient method to aggregate VoIP packets over ITTP protocol," *International Journal of Innovative Computing, Information and Control*, vol. 15, no. 3, pp. 1067–1077, 2019.
- [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 Standard for Information technology–Telecommunications and information exchange between systems Local and metropolitan area networks–Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, in *IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007)*, pp. 1–2793, 29 March 2012.
- [23] S. Yun, H. Kim, H. Lee and I. Kang, "100+ VoIP calls on 802.11b: 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] K. Mansour, J. Issam and E. Tahar, "Revisiting the IEEE 802.11n A-MPDU retransmission scheme," *Journal of IEEE Communications Letters*, vol. 23, no. 6, pp. 1097–1100, June 2019.
- [27] 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.
- [28] D. Jianhua, "An adaptive packet aggregation algorithm (AAM) for wireless networks," Ph.D. dissertation, Technological University Dublin, Ireland, 2013.
- [29] K. Hassine, "Performance study of future wireless networks IEEE 802.11 xy," Ph.D. dissertation, Higher School of Communication of Tunis, Tunisia, 2016.
- [30] R. Karmakar, C. Samiran and C. Sandip, "Impact of IEEE 802.11n/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.
- [31] 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.
- [32] E. Charfi, C. Gueguen, L. Chaari, B. Cousin and L. Kamoun, "Dynamic frame aggregation scheduler for multimedia applications in IEEE 802.11n networks," *Journal of Transactions on Emerging Telecommunications Technologies*, vol. 28, no. 2, pp. e2942, 2017.
- [33] S. Seytnazarov and Y. T. Kim, "QoS-aware adaptive a-MPDU aggregation scheduler for enhanced VoIP capacity over aggregation-enabled WLANs," in *IEEE/IFIP Network Operations and Management Sym.*, Taipei, Taiwan, pp. 1–7, 2018.
- [34] Z. D. Dudu, "Packet aggregation for voice over internet protocol on wireless mesh networks," Ph.D. dissertation, University of the Western Cape, South Africa, 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] T. Li, Q. Ni, D. Malone, D. Leith, Y. Xiao, T. Turletti *et al.*, "Aggregation with fragment retransmission for very high-speed WLANs," *Journal of IEEE/ACM Transactions on Networking*, vol. 17, no. 2, pp. 591–604, 2009.
- [37] A. Saif, M. Othman, S. Subramaniam and N. AbdulHamid, "Impact of aggregation headers on aggregating small A-MSDU in 802.11n WLANs," in *Proc. of Int. Conf. on Computer Applications and Industrial Electronics*, Kuala Lumpur, Malaysia, pp. 630–635, 2010.
- [38] A. Saif, M. Othman, S. Subramaniam and N. A. W. A. Hamid, "An enhanced A-MSDU frame aggregation scheme for 802.11n wireless networks," *Journal of Wireless Personal Communications*, vol. 66, no. 4, pp. 683–706, 2012.

- [39] A. Saif, M. Othman, S. Subramaniam and N. A. W. A. Hamid, "An optimized A-MSDU frame aggregation with subframe retransmission in IEEE 802.11n wireless networks," *Journal of Procedia Computer Science*, vol. 9, pp. 812–821, 2012.
- [40] N. Hajlaoui, I. Jabri, M. Taieb and M. Benjemaa, "A frame aggregation scheduler for QoS-sensitive applications in IEEE 802.11n WLANs," in *Proc. of Int. Conf. on Communications and Information Technology*, Hammamet, Tunisia, pp. 221–226, 2012.
- [41] N. Hajlaoui, I. Jabri and M. B. Jemaa, "Experimental performance evaluation and frame aggregation enhancement in IEEE 802.11n WLANs," *International Journal of Communication Networks and Information Security*, vol. 5, no. 1, pp. 48–58, 2013.
- [42] A. Saif and M. Othman, "A reliable A-MSDU frame aggregation scheme in 802.11n wireless networks," *Journal of Procedia Computer Science*, vol. 21, pp. 191–198, 2013.
- [43] A. Saif and M. Othman, "SRA-MSDU: Enhanced A-MSDU frame aggregation with selective retransmission in 802.11n wireless networks," *Journal of Network and Computer Applications*, vol. 36, no. 4, pp. 1219–1229, 2013.
- [44] B. Maqhat, M. D. Baba, R. A. Rahman and A. Saif, "Performance analysis of fair scheduler for A-MSDU aggregation in IEEE 802.11n wireless networks," in *Proc. of 2nd Int. Conf. on Electrical, Electronics and System Engineering*, Kuala Lumpur, Malaysia, pp. 60–65, 2014.
- [45] C. Shin, H. Park and H. M. Kwon, "Phy-supported frame aggregation for wireless local area networks," *Journal of IEEE Transactions on Mobile Computing*, vol. 13, no. 10, pp. 2369–2381, 2014.
- [46] H. Yeon, S. K. Sahu, W. Park, J. W. Lee and J. Lee, "Adaptive transmission scheme using dynamic aggregation and fragmentation in WLAN MAC," in *Proc. of IEEE Wireless Communications and Networking Conf.*, Istanbul, Turkey, pp. 1532–1537, 2014.
- [47] M. Moh, M. Teng-Sheng and K. Chan, "Error-sensitive adaptive frame aggregation in 802.11n WLAN," in *Proc. of Int. Conf. on Wired/Wireless Internet Communications*, Berlin, German, pp. 64–76, 2010.
- [48] F. Blanco, D. Wing, J. Navajas, M. Perumal and J. Saldana, "Tunneling Compressing and Multiplexing (TCM) Traffic Flows" Reference Model draft-saldana-tsvwg-tcmtf-08, Transport Area Working Group, 2015.
- [49] 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.
- [50] M. Abu-Alhaj, S. K. Manjur, R. Sureswaran, T. C. Wan, I. J. Mohamad *et al.*, "ITTP: A new transport protocol for VoIP applications," *International Journal of Innovative Computing, Information and Control*, Seoul, South Korea, vol. 8, no. 3, pp. 1–10, 2012.
- [51] S. Shin and H. Schulzrinne, "Measurement and analysis of the VoIP capacity in IEEE 802.11 WLAN," *Journal of IEEE Transactions on Mobile Computing*, vol. 8, no. 9, pp. 1265–1279, 2009.
- [52] H. P. Sze, C. L. Soung, J. Y. B. Lee and D. C. S. Yip, "A multiplexing scheme for h. 323 voice-over-IP applications," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 7, pp. 1360–1368, 2002.
- [53] B. Subbiah, S. Sengodan and J. Rajahalm, "RTP payload multiplexing between IP telephony gateways," in *Proc. of Global Telecommunications Conf.*, vol. 2, pp. 1121–1127, 1999.
- [54] K. Kim, S. Ganguly, R. Izmailov and S. Hong, "On packet aggregation mechanisms for improving VoIP quality in mesh networks," in *Proc. of IEEE 63rd Vehicular Technology Conf.*, vol. 2, pp. 891–895, 2006.
- [55] S. Jung, S. Hong and P. Park, "Effect of robust header compression (ROHC) and packet aggregation on multi-hop wireless mesh networks," in *Proc. of the Sixth IEEE Int. Conf. on Computer and Information Technology*, pp. 91, 2006.