electronics

Article

Performance of WLAN in Downlink MU-MIMO Channel with the Least Cost in Terms of Increased Delay

Lemlem Kassa [1,*], Jianhua Deng [1], Mark Davis [2] and Jingye Cai [1]

[1] School of Information and Software Engineering, University of Electronic Science and Technology China (UESTC), Chengdu 610054, China
[2] Communication Network Research Institute (CNRI), Technological University Dublin, D08 NF82 Dublin, Ireland
* Correspondence: lemlem.kassa@aastu.edu.et

Abstract: To improve the performance of IEEE 802.11 wireless local area (WLAN) networks, different frame-aggregation algorithms are proposed by IEEE 802.11n/ac standards to improve the throughput performance of WLANs. However, this improvement will also have a related cost in terms of increasing delay. The traffic load generated by mixed types of applications in current modern networks demands different network performance requirements in terms of maintaining some form of an optimal trade-off between maximizing throughput and minimizing delay. However, the majority of existing researchers have only attempted to optimize either one (to maximize throughput or minimize the delay). Both the performance of throughput and delay can be affected by several factors such as a heterogeneous traffic pattern, target aggregate frame size, channel condition, competing stations, etc. However, under the effect of uncertain conditions of heterogeneous traffic patterns and channel conditions in a network, determining the optimal target aggregate frame size is a significant approach that can be controlled to manage both throughput and delay. The main contribution of this study was to propose an adaptive aggregation algorithm that allows an adaptive optimal trade-off between maximizing system throughput and minimizing system delay in the WLAN downlink MU-MIMO channel. The proposed approach adopted different aggregation policies to adaptively select the optimal aggregation policy that allowed for achieving maximum system throughput by minimizing delay. Both queue delay and transmission delay, which have a significant impact when frame-aggregation algorithms are adopted, were considered. Different test case scenarios were considered such as channel error, traffic pattern, and number of competing stations. Through system-level simulation, the performance of the proposed approach was validated over the FIFO aggregation algorithm and earlier adaptive aggregation approaches, which only focused on achieving maximum throughput at the expense of delay. The performance of the proposed approach was evaluated under the effects of heterogenous traffic patterns for VoIP and video traffic applications, channel conditions, and number of STAs for WLAN downlink MU-MIMO channels.

Keywords: adaptive frame aggregation; downlink MU-MIMO; wireless local area network (WLAN); network traffic; queue delay; throughput; transmission delay

Citation: Kassa, L.; Deng, J.; Davis, M.; Cai, J. Performance of WLAN in Downlink MU-MIMO Channel with the Least Cost in Terms of Increased Delay. *Electronics* **2022**, *11*, 2851. https://doi.org/10.3390/electronics11182851

Academic Editors: Andrey Lyakhov and Raed A. Abd-Alhameed

Received: 13 June 2022
Accepted: 26 August 2022
Published: 9 September 2022

1. Introduction

Due to the advancement of wireless technologies, IEEE 802.11-based networks are becoming more popular and different technologies have been introduced to improve throughput performance. Multiuser, multiple-input, multiple-output (MU-MIMO) is among the technologies at the physical layer introduced in the IEEE 802.11ac standard to accommodate the increasing demand for high data-transmission rates by allowing a single access point (AP) that supports simultaneous transmission for up to a maximum of eight users at a time [1,2]. This is one of the most crucial technologies that has driven wireless local area networks (WLANs) into the gigabit era. Moreover, the wireless medium has a high overhead in terms of bytes that can be higher than the actual payload. To amortize these overheads, which include the medium access control (MAC) and physical (PHY) headers, acknowledgments (ACK), backoff time, and interframe spacing, the standard also introduced a frame-aggregation scheme that has contributed to high data throughput by combining multiple frames, also known as MAC service data units (MSDUs), into a single transmission unit [1]. The performance of WLAN depends on different performance factors at different layers of the network protocol stack, such as at the PHY and MAC layers. For instance, the frequency channel, modulation and coding schemes, transmitter power, etc., at the PHY layer; and the retry limit, frame size, contention window size, maximum number of backoffs, etc., at the MAC layer have a significant impact on the performance of WLAN. Optimizing these parameters would improve the system performance of WLAN. If a wireless frame size is large, a bit error can destroy the whole frame, thus the frame success rate decreases, and thus the throughput performance is degraded [3]. On the other hand, if the frame size is shorter, the overhead frames such as MAC and PHY headers occupy a large portion of the transmitted frame and thus degrade the transmission efficiency [3].

The IEEE 802.11 standard specifies a constant-length aggregation strategy regardless of the traffic pattern and channel conditions in the actual network. This contributes to the reduction in channel access overhead. However, utilizing the maximum aggregation size may not be optimal in all situations because it may lead to an increase in the delivery of error frames and retransmissions [4]. However, both throughput and delay are the most important performance metrics that need to be considered in frame-aggregation algorithms [5]. Frame aggregation allows protocol overhead to be reduced, thus it can significantly improve throughput performance, but this improvement will also have a related cost in terms of increasing the delay [5–7]. The time for which more frames wait before transmission in a buffer is the main factor of delay when using frame-aggregation algorithms [5–9]. Therefore, this indicates that there is a trade-off between increasing throughput and increasing delay. Network delay is an important performance characteristic of the IEEE 802.11 wireless network, and it is usually categorized into different parts such as processing delay, transmission delay, queue delay, and propagation delay [5,7]. In this study, queue delay and transmission delay were considered. Queue delay is defined as the queuing time of the first arrival packet waiting in the buffer [7,8]. In this study, it was computed by considering the arrival time of the first frame (i.e., the time from when it arrived at the AP's buffer until the time it was begun to be transmitted) [7,8]. Transmission delay is defined as the time when a station begins to check the channel state (i.e., idle or busy) for transmitting a frame until it receives an ACK of the frame [6]. Average minimum delay is defined as the average delay under saturation conditions [5,6]. Under the conditions of an unpredictable heterogeneous traffic pattern and channel conditions, different aggregation algorithms provide different performances in terms of throughput and delay. In addressing these challenges, this study proposed to realize the optimal system throughput of WLAN in a downlink MU-MIMO channel for minimizing system delay by employing a dynamic adaptive aggregation selection scheme.

Therefore, the main contribution of this paper was to propose an adaptive aggregation algorithm that attempted to achieve an optimal trade-off between maximizing throughput and minimizing delay. The proposed approach adopted different aggregation policies [10] to adaptively select the optimal aggregation policy that maximized the system throughput while minimizing the cost of delay. Different test-case scenarios were considered such as channel error, traffic pattern, and number of competing stations. Through system-level simulation, the performance of the proposed approach was validated over the FIFO aggregation algorithm and earlier adaptive aggregation approaches [11], which do not consider the issue of delay under the effects of heterogenous traffic patterns, channel conditions, and number of STAs for WLAN downlink MU-MIMO channels. To the best of our knowledge, this was the first attempt to exploit the trade-off between both maximizing throughputs and minimizing delay while considering the impact of queue delay and transmission delay in the WLAN downlink MU-MIMO channel. Moreover, this study addressed the challenges of heterogeneous traffic demand and channel conditions while considering a traffic mix of VoIP and video.

The rest of the paper is organized as follows. In Section 2, we introduce related works on frame-aggregation schemes and the performance challenges of multiuser transmissions in the WLAN downlink MU-MIMO channel. A detailed problem description of the proposed approach is given in Section 3. In Section 4, the results and discussions are presented to evaluate the performance of the proposed approach under various channel conditions, traffic models, and number of stations. Finally, the conclusions are given in Section 5.

2. Related Work and Our Motivation

In this section, some previous works and the effects of approaches to frame size determination on the performance of WLAN are discussed, mainly focusing on the downlink MU-MIMO channel.

2.1. Related Work

The majority of existing researchers have only attempted to achieve optimization via the trade-off between maximizing throughput and minimizing delay. However, under the effects of uncertain conditions of heterogeneous traffic patterns and channel conditions in a network, throughput and delay are the two most important performance metrics [5–9] that must be analyzed in frame-aggregation algorithms. Throughput can be improved by increasing the size of the average payload or by reducing the size of overhead frames such as the MAC and PHY headers in a frame. This can be achieved by adopting frame-aggregation algorithms. An adaptive aggregation algorithm called the adaptive aggregation mechanism (AAM) was proposed in [5] that allowed an adaptive trade-off between maximizing throughput and minimizing delay in a single-user WLAN channel. This approach considered the varying nature of the packet size and the packet arrival time to assemble the target aggregate packet size within the minimum delay. However, this approach did not consider the effects of transmission errors and competing stations. The work in [6] demonstrated the existence of the throughput upper limit (TUL) in a WLAN while considering assumptions such as only one sender and one receiver existing in the wireless network, the sender always having frames to be transmitted, and each frame having the same size operating in the distributed coordination function (DCF) mode. They also assumed ideal channel conditions when there were no transmission errors present. The delay lower limit (DLL) in the IEEE 802.11 wireless network using the DCF model was also demonstrated in [6]. The researchers considered best-case scenarios such as an ideal channel condition with only one active station that always has frames to send while other stations can only receive frames and send acknowledgments (ACKs) during any transmission cycle. The main source of delay in the frame-aggregation algorithm was the time spent waiting for more frames to arrive. However, the assumptions of this approach were unrealistic and could not function in the MU-MIMO channel. Moreover, this approach only considered the transmission delay. However, delay of both the queuing time and waiting time in the MAC are the key factors of delay that can significantly increase when using frame aggregation [5]. In [9], a novel method was proposed to determine the frame-aggregation size in a MU-MIMO channel to improve channel utilization while considering the delay during which data frames waited in transmission queues. They attempted to reduce the delay by appropriately determining the aggregation size according to the traffic variation. However, the main focus of this study was to enhance channel utilization, and the effects of channel errors were not elaborated.

Some studies have contributed frame-size-aggregation schemes in WLAN downlink MU-MIMO channels [10–20]. For instance, the algorithm in [10] proposed a new approach that aimed to enhance the system throughput performance of a WLAN by employing a dynamic adaptive aggregation selection scheme to determine the optimal length of the frame size in downlink MU-MIMO transmission. The effects of heterogeneous traffic demand among spatial streams were considered under the assumption of an ideal channel. According to the simulation results, the maximum system throughput performance and channel utilization were achieved. By extending this work, an adaptive frame-aggregation algorithm was proposed in [11] while considering the effects of transmission errors. However, both of these studies did not consider the expense of delay. Thus, this led to a suboptimal solution. The current study enhanced the scheme of [11] by considering the issue of delay to achieve the maximum system throughput performance of a WLAN in the downlink MU-MIMO channel.

The frame-size-optimization problem has been studied by several researchers for IEEE 802.11 networks. By employing a specific procedure of dynamically adjusting the frame size, [12] proposed a method that dealt with frame-size estimation based on the extended Kalman filter for saturated networks. They derived the mathematical equation for throughput, which was a function of the frame size. The optimal frame size was obtained using differential calculus. Likewise, Bianchi's Markov chain model studied the relationship between the throughput and frame size in IEEE 802.11 WLANs [13]. However, the assumption of this work was an ideal channel, which was unrealistic. According to the simulation results, the throughput increased with the frame size; i.e., the larger the frame size, the better the throughput. However, the cost of delay increased when increasing the throughput [5], thus the results provided a suboptimal solution. A machine-learning-based frame-size-optimization approach that considered channel conditions and contention effects of users was proposed in [14] by extending Bianchi's model [13] as the main simulation environment. According to the simulation results, the frame-size optimization was effectively achieved to maximize the throughput performance of the WLAN at the expense of delay. This work did not consider frame aggregation but limited its contribution to dynamic frame fragmentation and defragmentation to maintain backward compatibility with IEEE 802.11a/b/g WLANs. An adaptive algorithm for frame-size optimization was proposed in [15] that allowed an ARQ protocol to dynamically optimize the packet size based on estimates of the channel bit errors. The main strategy of this study was to make estimates of the channel bit-error rate; the researchers considered the acknowledgment history, and based on that, the optimal packet size could be determined. However, this approach is not suitable for IEEE 802.11 WLAN environments. In general, all of the above studies mainly focused on maximizing throughput at the expense of delay. However, there is always a trade-off between maximizing throughput and minimizing delay when adopting frame-size-determination strategies [5].

A data-frame-construction scheme called DFSC was proposed in [16] to determine the length of a multiuser (MU) frame with the aim to maximize the transmission efficiency by considering the status of buffers and transmission bit rates of stations in both uplink and downlink multiuser transmissions. However, they did not consider the effects of channel errors, which can reduce the transmission performance due to excessive retransmissions of frames received in error, and the cost of this delay was not elaborated upon. A frame-size-based aggregation scheme was proposed in [17]; the authors demonstrated that both the queueing length and number of active nodes had significant impacts on the system throughput performance. The main approach of this paper was to generate the same frame length in all spatial streams that could maximize the system throughput performance. However, this study focused on maximizing throughput while ignoring the cost of delay.

Some works in the literature focused on the padding problem. According to [18,19], the authors improved the transmission efficiency in the downlink MU-MIMO channel by replacing padding bits with data frames from other users in one stream to fill the space of frame padding, violating the rules of MU transmissions. However, these approaches increased the complexity of both the transmission and reception process in wireless communication, which requires modification of the standard to allow the transmission to multiple destinations within a special stream. A frame-duration-based frame-aggregation scheme was proposed in [20] that employed criteria for selecting a receiving mobile terminal (MT). This approach provided high priority to the MT expecting high throughput in the next MU-MIMO transmission and having a large amount of data while reducing the signaling overhead. By equalizing the transmission time of all spatial streams in all MTs according to their modulation and coding (MCS) level, the authors achieved maximum performance of system throughput and minimized the space channel time in the WLAN's downlink MU-MIMO channel. However, the frame-size-determination scheme in this study did not consider the expense of delay. Although all the above proposals contributed several schemes to enhance the performance of WLANs, none of them proposed the cost of delay. To the best of our knowledge, there is little research that explored a realization of the optimal trade-off between maximizing throughput and minimizing delay in WLAN downlink MU-MIMO transmission.

2.2. Motivation for This Work

The earlier [10,11] dynamic adaptive frame-aggregation selection schemes could maximize the system throughput performance of a WLAN in terms of the maximum system throughput and channel utilization and the minimum space channel time. However, this approach does not consider the issue of delay. The motivation of this work was an intention to extend the previous work [11] to contribute an adaptive aggregation algorithm aiming to achieve the optimal trade-off between maximizing system throughput and minimizing system delay.

3. Proposed Approach

This section proposes a frame-aggregation scheme that allows an adaptive trade-off between maximizing throughput and minimizing delay by employing a dynamic adaptive aggregation selection mechanism for the WLAN downlink MU-MIMO channel. In this study, the challenges of heterogenous traffic patterns, channel errors, and the number of stations, which severely affect the performances of both throughput and delay, were considered. Any improvement achieved by adopting aggregation experiences an increasing delay. Therefore, the adaptive aggregation mechanism is one of the methods to control the performance of throughput and delay.

This study proposed to enhance the performance of the adaptive aggregation algorithm in [11] by considering the issues of queue delay and transmission delay, which have a significant impact when frame-aggregation algorithms are adopted. The main contribution of this study was to improve the system throughput of WLAN with the least cost in terms of a delay increase. Generating traffic was the first operation achieved by using Preto, Weibull, and fBM traffic models adopted from [10] for the MAC service data unit (MSDU) data frames in bytes. According to our assumption, constant frame sizes of 100 bytes for a traffic mix of 75% VoIP and 1000 bytes for 25% video traffic [10] were considered. All user frames were buffered in the AP, including the new arrival frame and remaining frames that were not selected for aggregation. Remaining frames could occur due to the target aggregate frame size determined by the type of aggregation policy employed and frames that remained in the buffer due to the maximum A-MPDU frame-size limitation specified in the IEEE 801.11ac standard [10]. AP buffer frames as large as the maximum storage capacity of 50 MB were allowed. However, if the AP buffer was full, the reception of new frames was denied, and the AP continued with the output process until some space became free at the buffer.

Different aggregation algorithms experience different performances in terms of throughput and delay mainly due to unpredictable traffic patterns and channel conditions in the actual network. Thus, the different aggregation policies proposed in [10] experience different performance in terms of both throughput and delay. For instance, FIFO FA allows a maximum frame-size aggregation as long as the maximum aggregate frame size allowed by IEEE 802.11, which improves the throughput. Moreover, it reduces the queue delay of frames before transmission. However, a longer frame size will incur a longer transmission delay. Equal Frame Size FA allows all streams to have an equal aggregation size while considering the station's queue size. This approach always selects the smaller queue size to assemble the target aggregate frame size. Therefore, if the traffic is highly bursty; i.e., if the queue size of the stations has a high variation, the aggregation overhead will be higher due to shorter payload frame aggregation. However, on the contrary, the shorter frame-size aggregation could perform better in error-prone channel conditions [11]. The Equal MPDUs Agg FA allows an equal number of aggregated frames to be assembled in all streams. Similarly, this aggregation policy is affected by the traffic pattern; i.e., if the traffic is highly bursty, the target aggregate frame size will be shorter, which degrades the throughput and likewise increases the queue delay due to remaining frames left in the buffer before transmission. The Avg Num MPDUs FA considers the average number of frames among streams to assemble the target aggregate frame size. This approach can perform better when the traffic variation among streams is bursty, but it also experiences a delay due to the remaining frames in the buffer. Therefore, this indicates that a trade-off exists between maximizing throughput and minimizing delay when adopting frame-aggregation algorithms due to traffic patterns and channel condition. This study addressed this trade-off by proposing an adaptive aggregation algorithm that allowed a dynamic adaptive aggregation selection scheme by adopting different aggregation policies (FIFO FA, Equal Frame Size FA, Equal MPDUs Agg FA, and Avg Num MPDUs FA) [10]. The aim was to realize an optimal trade-off between maximizing throughput and minimizing delay. Figure 1 shows a flowchart of the proposed adaptive aggregation algorithm while considering the delay.

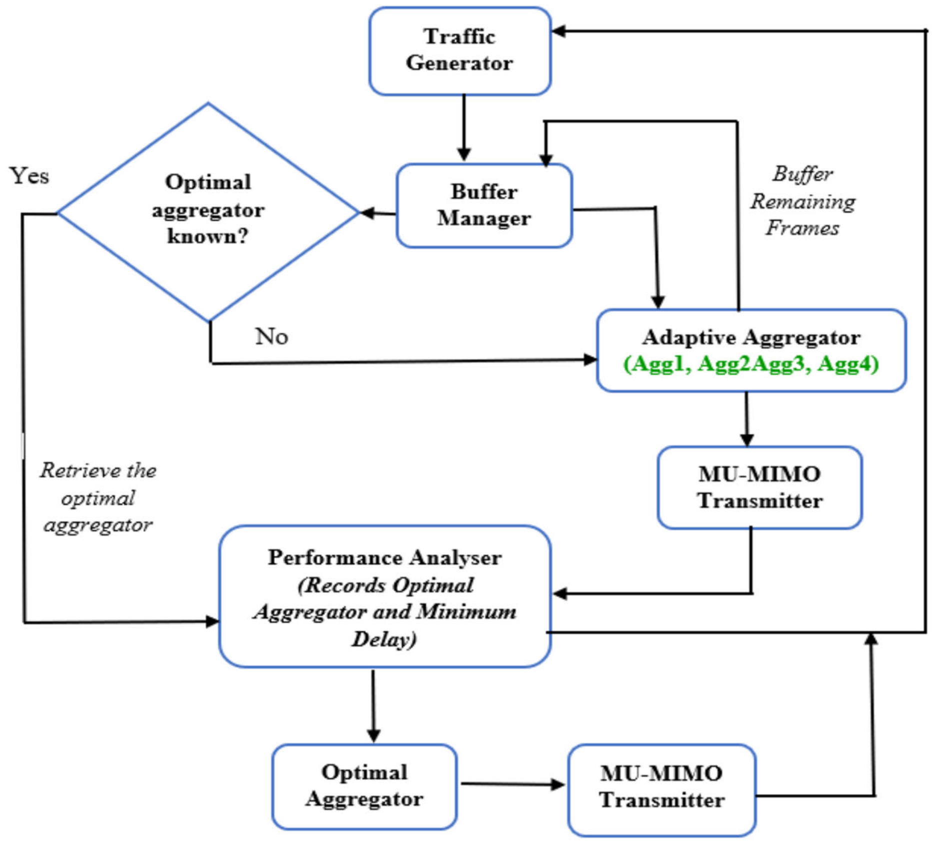


Figure 1. Proposed adaptive aggregation algorithm.

As shown in Figure 1, once the traffic was generated and buffered, the AP performed the optimal aggregation policy selection under the different channel conditions (SNR = 5, 8, or 20 dB), traffic models (Pareto, Weibull, or fBM), and number of STAs while considering a traffic mix of 75% VoIP and 25% video. Then, the aggregation manager aggregated the data frame of each user while employing the four different aggregation policies depicted in [10], such as Agg1, Agg2, Agg3, and Agg4 in Figure 1, which represented the FIFO FA (baseline approach), Equal Frame Size FA, Equal MPDUs Agg FA, and Avg Num MPDUs FA aggregation policies, respectively. Then, the aggregated data frames of each aggregation policy were transmitted one by one to the receiving STAs (receivers) using the

MU-MIMO technique. The receiving stations received different aggregated frame sizes that were obtained using the different aggregation policies. Then, the performance analyzer in the AP evaluated the performance of each aggregation policy, determined which one provided the minimum system delay, and recorded the optimal aggregation policy and the corresponding minimum delay. This record was then utilized in future operations when a similar channel condition and traffic pattern and number of stations occurred in the system. Therefore, from all possible frame sizes obtained from the different aggregation policies, the optimal system frame size was the one obtained from the optimal aggregation policy with the minimum delay. However, if the AP determined the optimal aggregation policy with the given traffic pattern, channel condition, and number of competing stations by retrieving the performance analyzer, frame aggregation could be constructed by the optimal aggregator and the MU-MIMO transmission could be made promptly, as shown in the flowchart.

In this study, delay was defined as the average time to successfully transmit a data frame from the source station (AP) to the destination station, while the minimum average delay was the average delay under saturation conditions in the wireless network. As discussed, the aggregation process began after the arrival of a certain numbered frame in the buffer. Moreover, the different aggregation policies adopted different aggregation strategies when selecting frames for aggregation (FIFO FA (baseline approach), Equal Frame Size FA, Equal MPDUs Agg FA, and Avg Num MPDUs FA). Thus, in this study, we considered the queue delay while considering both the waiting delay and aggregation delay. Queue delay (*QDelay*) was defined as the sum of time spent waiting for more packets to arrive (waiting delay) and waiting time for aggregation (aggregation delay), which was the amount of time frames waited in the buffer when they were not selected for aggregation. The frame-aggregation delay was the amount of time between the arrival time of the first frame in the buffer until the time it could start to be transmitted [8]. Equations (1)–(3) illustrate how the queue delay was computed in this study. Both the waiting delay and aggregation delay in the equations considered the time difference between the arrival time of the old frame and just-arrived frames for n number of stations in the downlink MU-MIMO channel streams. The maximum number of stations assumed in this study was 4.

$$Waiting\ Delay_t = \sum_{i=1}^{n} (ArrTime_{lastP}\ - ArrTime_{firstP}) \tag{1}$$

$$Aggregation\ Delay_t = \sum_{i=1}^{n} (AggDelay_{lastP}\ - AggDelay_{firstP}) \tag{2}$$

Thus, *QDelay* for each aggregation policy such as FIFO FA (baseline approach), Equal Frame Size FA, Equal MPDUs Agg FA, and Avg Num MPDUs FA was defined as shown in Equation (3):

$$QDelay_t = \sum (WaitingDelay_t + AggregtionDelay_t) \tag{3}$$

where:

- n is number of stations;
- t is the tth simulation time;
- $ArrTime_{lastP}$ is the arrival time of the last frame;
- $ArTime_{firstP}$ is the arrival time of the first frame;
- $AggDelay_{lastP}$ is the aggregation delay of the last frame;
- $AggDelay_{firstP}$ is the aggregation delay of the first frame.

The transmission delay ($TDelay$) was defined as the time when a station began to check the channel state (i.e., idle or busy) for transmitting a frame until it received an ACK of the frame [6]. The time required for completing a single MU-MIMO transmission by each of the aggregation policies (FIFO FA (baseline approach), Equal Frame Size FA, Equal

MPDUs Agg FA, and Avg Num MPDUs FA) was determined by the transmission time of the longer frame, which was defined as $max(TData_i)$ [11], as shown in Equation (4):

$$TDelay_t = TDIFS + BOTime + max(TData_i) + NumSTA(TSIFS + TBA) + \dots \quad (4)$$

Therefore, the average system delay for a single MU-MIMO transmission in different aggregation policies (FIFO FA (baseline approach), Equal Frame Size FA, Equal MPDUs Agg FA, and Avg Num MPDUs FA) was defined as the ratio of the sum of the queue delay and transmission delay and the number of stations, as illustrated in Equation (5):

$$AverageSystemDelay_t = \frac{(\sum_{i=1}^{n} QDelay_t + TDelay_t)}{n} \quad (5)$$

where n is number of stations in the network and t is the tth simulation time.

Finally, the optimal aggregation policy with the minimum system delay was obtained by comparing the different aggregation algorithms (FIFO FA (baseline approach), Equal Frame Size FA, Equal MPDUs Agg FA, and Avg Num MPDUs FA) using the performance analyzer, as illustrated in Equation (6):

$$\begin{aligned} &[SysDelayMin_t, Index\ \_Agg\_Policy_t] \\ &= \min(SysDelayAgg1, SysDelayAgg2, SysDelayAgg3, SysDelayAgg4) \end{aligned} \quad (6)$$

Thus, $SysDelayMin_t$ gave the minimum system delay and the index of the optimal aggregation policy $Index\ \_Agg\_Policy_t$ at time t. It was assumed that the aggregation policies were arranged in the order of 1 to 4, respectively, for Agg1, Agg2, Agg3, and Agg4. Therefore, if the $Index\ \_Agg\_Policy_t = 1$, the minimum system delay was achieved by Agg1. Otherwise, if the $Index\ \_Agg\_Policy_t = 4$, it was Agg4. The system throughput was the average data rate at which the AP could successfully deliver to all receiving stations [10,11]. It was computed as the ratio of the sum of all the successful frame sizes of the system over the total channel transmission time.

4. Results and Discussion

In this section, we evaluate the performance of the proposed adaptive approach to maximize the system throughput of the WLAN downlink MU-MIMO channel with the least cost in terms of a delay increase by considering the effects of channel conditions, heterogeneous traffic patterns, and number of stations. The simulation was conducted based on the IEEE 802.11ac standard for the MAC and PHY layer specifications. Table 1 shows the detailed simulation parameter settings considered in the experiment.

Table 1. Simulation parameters.

Parameters	Symbol	Value
Number of antennas at AP	N_{Ant}	4
Number of stations	Num_{STA}	2–4
Average data frame length	LData	100 bytes for VoIP, 1000 bytes for video
Traffic rate		10 Kbps for VoIP, 100 Mbps for video
Data rate		260 Mbps per STAs
Minimum window size	CW-min	15
Transmission time slots	Tslot	16 microseconds
Average A-MSDU length		11,454 Byte

Table 1. *Cont.*

Parameters	Symbol	Value
Max. number of MPDU frames aggregated		64
Max. A-MPDU length		1.0 Mbyte
SNR		5, 8, 20 dB
Max. number of retransmissions	N_{Ret}	3
Buffer size at the AP	BufAP	50 MB

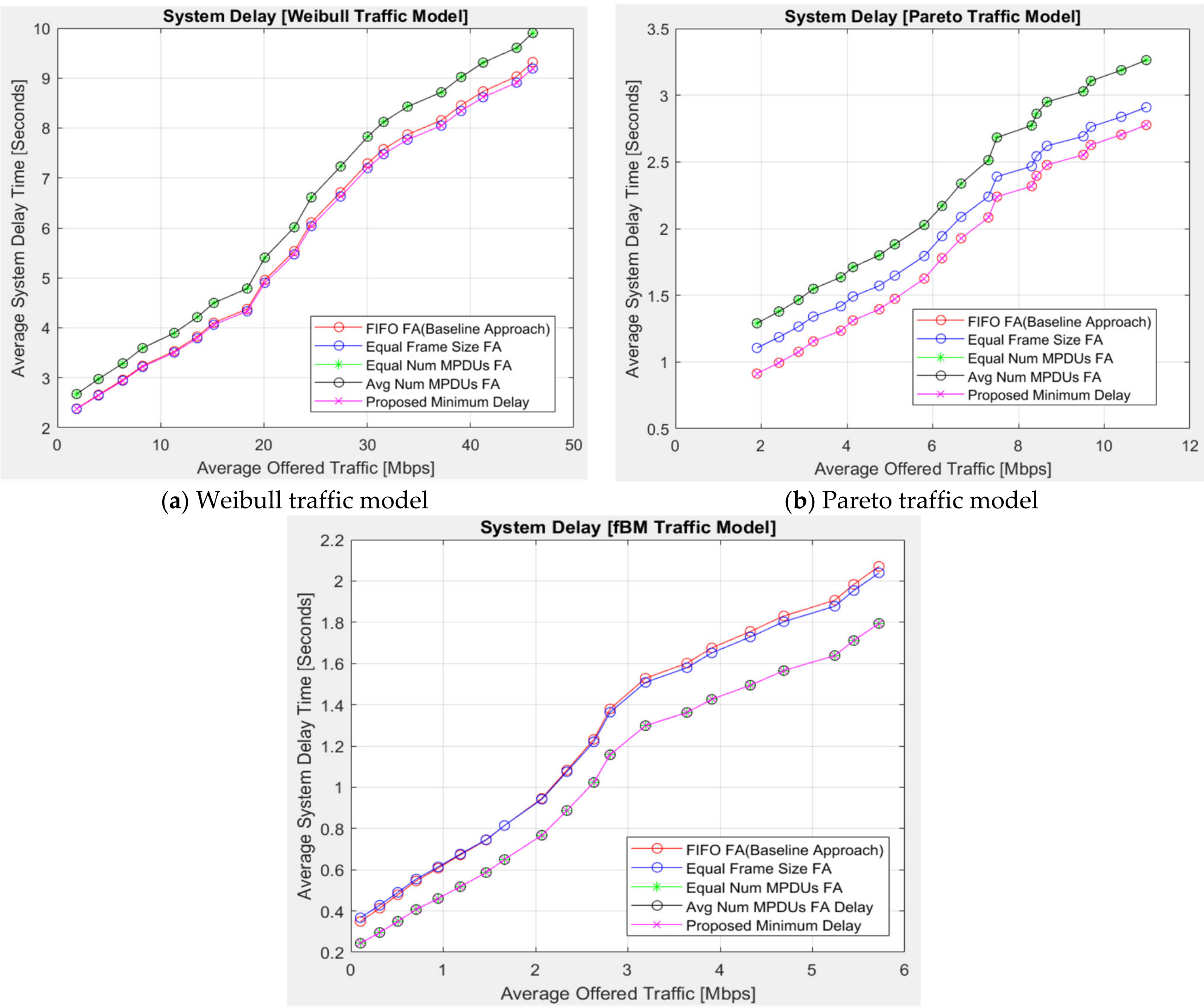
4.1. Experimental Procedure

The performance of the proposed approach was evaluated over the FIFO FA (baseline approach) and Adaptive FA Conv. approach [11] in terms of the system throughput performance while considering the minimum delay. FIFO FA (baseline approach) does not employ an adaptive aggregation scheme; the Adaptive FA Conv. approach [11] is an adaptive aggregation algorithm that considers channel errors but ignores the expense of delay, aiming to achieve maximum throughput only. The acronyms 'FA' and 'Conv.' in this paper refer to 'frame aggregation' and 'conventional', respectively. In Section 4.2, we will evaluate the cost of delay experienced in different aggregation algorithms and the performance of the proposed approach achieved in terms of the minimum delay. The proposed adaptive approach will be evaluated under the effects of different traffic models such as Pareto, Weibull, and fBM in Section 4.3. Then, the performance of the proposed approach under the effect of channel conditions while considering SNR = 5, 8, and 20 dB is evaluated in Section 4.4. Finally, the performance of the proposed approach under a varying number of STAs (two, three, or four) is evaluated in Section 4.5. All experiments were conducted with a traffic mix of 75% VoIP and 25% video with a constant frame size of 100 bytes or 1000 bytes and a data rate of 10 kbps or 100 Mbps, respectively.

4.2. Performance of System Delay in Different Aggregation Algorithms

This experiment demonstrated the performance of delay in different aggregation algorithms and the performance of the proposed approach achieved by employing a dynamic adaptive aggregation selection scheme for the minimum delay. Due to the effects of heterogenous traffic demand and channel conditions in the actual network, it was difficult to achieve the maximum throughput while minimizing delay by employing a specific frame-aggregation algorithm. In addressing these challenges, our proposed approach attempted to achieve an optimal trade-off between maximizing throughput and minimizing delay by employing an adaptive aggregation selection scheme. The results displayed in Figure 2 show the performance of the delay in different traffic models when the average offered traffic increased. As the results showed, when the traffic load increased, the system delay time increased similarly in all traffic scenarios. However, due to the adaptive aggregation policy selection scheme employed in the proposed approach, it always attempted to achieve the minimum delay.

According to the results shown in Figure 2, the Equal Frame Size FA contributed the minimum delay of 9 s when the traffic load increased in the Weibull traffic model; however, the Equal Num MPDUs FA and Avg Num MPDUs FA achieved the maximum delay of 10 s. Therefore, the proposed approach achieved the minimum delay because the proposed approach adopted a dynamic aggregation selection strategy with the minimum delay. In the Pareto traffic model, the FIFO FA (baseline approach) contributed a minimum delay of 2.7 s as compared to the maximum delay of 3.3 s achieved by the Avg Num MPDUs FA and the Equal Num MPDUs FA. Moreover, the proposed approach achieved a better performance in the Weibull and fBM traffic models than in the Pareto traffic model because the traffic rate among streams was highly bursty in the Pareto traffic model. Therefore, under such a type of traffic pattern, the FIFO FA (baseline approach) outperformed because the number of frames waiting in the buffer was higher than when using the Equal Frame size FA, Avg Num MPDUs FA, and Equal Num MPDUs FA. On the contrary, the results showed that the fBM traffic model had the shortest system delay of 1.7 s contributed by the Equal MPDUs Agg FA and Avg Num MPDUs FA when the average traffic load increased, whereas the FIFO FA (baseline approach) achieved the maximum 2 s delay. In general, the results illustrated that a minimum delay could be achieved by employing an adaptive aggregation selection strategy.



(**a**) Weibull traffic model (**b**) Pareto traffic model

(**c**) fBM traffic model

Figure 2. Performance of average system delay time under the effects of heterogenous traffic models when SNR = 8 dB and $\mathrm{Num}_{\mathrm{STA}} = 4$.

4.3. Performance under the Effect of Various Traffic Models

In this experiment, the proposed approach was evaluated under the effects of different traffic models such as Pareto, Weibull, and fBM [10]; SNR = 8 dB; and $\mathrm{Num}_{\mathrm{STA}} = 4$. This experiment demonstrated how heterogeneous traffic patterns affected the optimal throughput performance with the minimum delay in the WLAN downlink channel. The FIFO FA (baseline approach), Equal Frame Size FA, Equal MPDUs Agg FA, Avg Num MPDUs FA, and Adaptive FA Conv. approach [11] were considered to evaluate the performance of the proposed approach in terms of achieving the maximum system throughput under the conditions of different traffic models.

As the results in Figure 3 show, due to the effects of heterogenous traffic patterns among streams, different system throughput performances were achieved by different traffic models. The proposed approach achieved the maximum performance of 583 Mbps using the Weibull traffic model and 504 Mbps using the fBM traffic model compared to the FIFO FA (baseline approach), which approached the maximum throughput performance of the Adaptive FA Conv. approach, which only considered the maximum throughput with the cost of maximum delay. On the contrary, the proposed approach achieved the minimum system throughput of 279 Mbps using the Pareto traffic model. This was because the highest bursty traffic behavior in the Pareto traffic allowed the waiting time of frames in a queue before transmission to maximize in terms of queue delay. Thus, the FIFO FA (baseline approach) contributed to the optimal system throughput with minimum system delay. This indicated that the performance of the proposed approach was better at coping with the less bursty traffic scenarios such as the Weibull and fBM traffic models. In general, these results indicated that traffic patterns in the network determined the system performance. The results also demonstrated that the proposed adaptive approach achieved a comparable result with the maximum system throughput achieved by the Adaptive FA Conv. approach, which was particularly focused on achieving the maximum system throughput at the expense of delay, particularly in the Weibull and fBM traffic models.

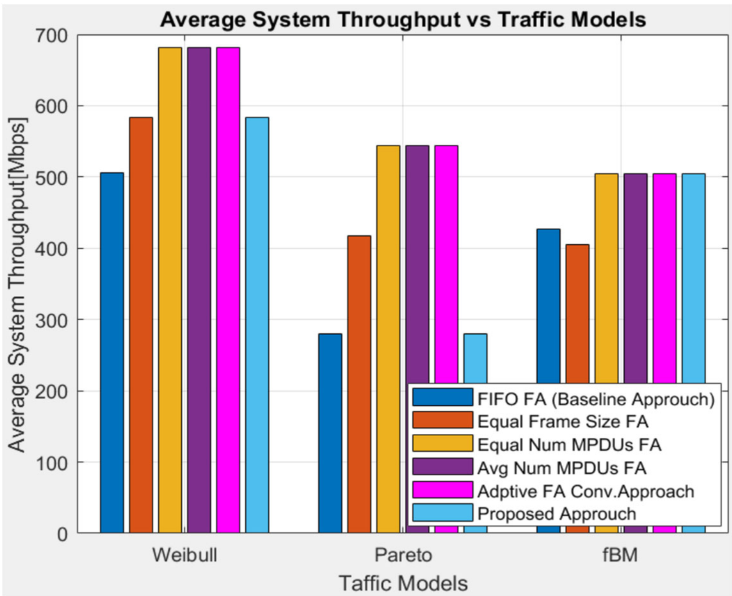
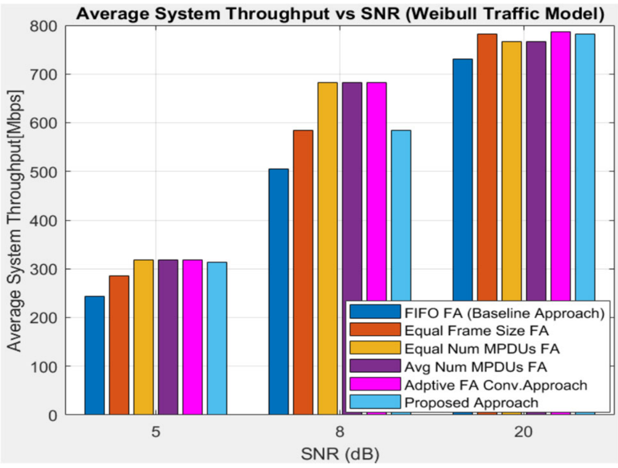


Figure 3. Performance of average system throughput under the effects of heterogenous traffic models when SNR = 8 dB.

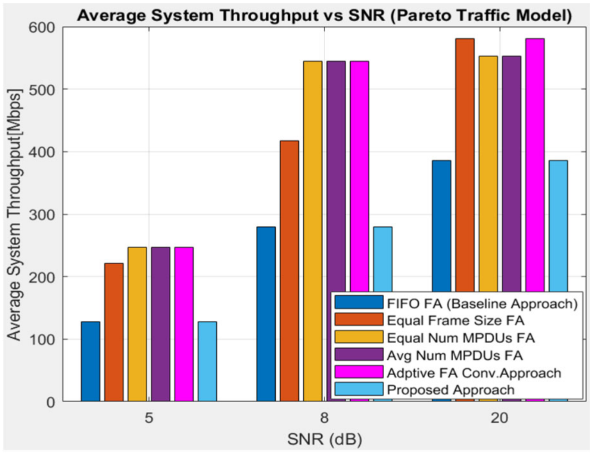
4.4. Performance under the Effects of Channel Conditions

The performance of the proposed approach under different channel conditions when SNR = 5, 8, and 20 dB and Num_{STA} = 4 was evaluated as shown in Figure 4a–c for the cases of different traffic models such as Pareto, Weibull, and fBM. According to the results, the system throughput performance increased when the channel quality improved from 5 dB to 20 dB in all traffic models. In this regard, the proposed approach achieved the lowest performance of 128 Mbps in the Pareto traffic model, as shown in Figure 4b, and the maximum of 312 Mbps was achieved by the Weibull traffic model when the channel condition was the worst; i.e., SNR = 5 dB. On the contrary, under the near-ideal channel conditions; e.g., with an SNR of 20 dB as shown in the figure, the system throughput performance was almost optimal in all approaches due to the lower frame error rate that occurred under the near-ideal channel condition, and the queue delay due to retransmitted frames was less. However, the proposed approach achieved the maximum performance of 781 Mbps using the Weibull traffic model, 582 Mbps using the fBM traffic model, and

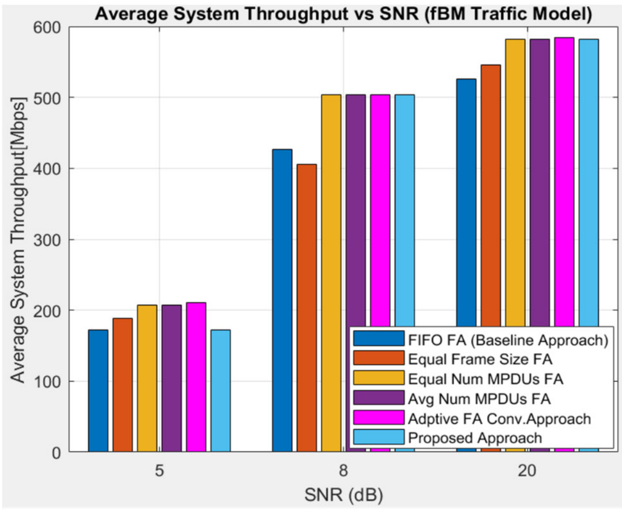
385 Mbps performance using the Pareto traffic model. This indicated that the proposed approach provided a good performance in the less bursty traffic model of the Weibull and fBM traffic models than in the bursty traffic model of Pareto traffic [10]. On the contrary, the Adaptive FA Conv. approach always achieved the maximum performance because it did not consider the cost of delay. However, the performance of the proposed approach was better because it adopted a dynamic adaptive aggregation selection scheme in terms of maximum throughput with the least cost of increasing delay.



(**a**) Weibull traffic model



(**b**) Pareto traffic model



(**c**) fBM traffic model

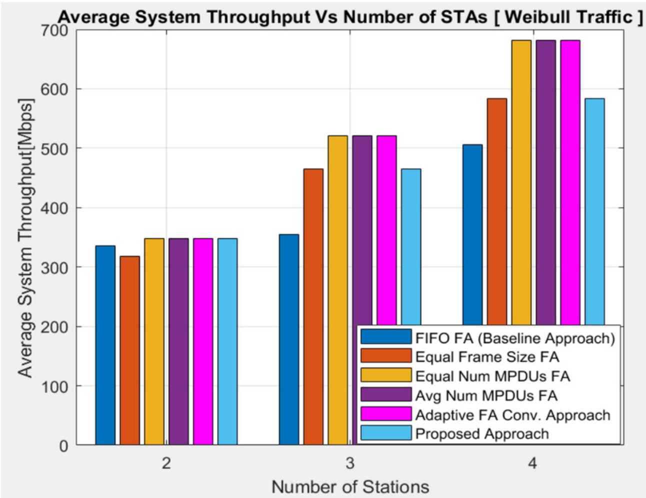
Figure 4. System throughput versus SNR for different traffic models such as Pareto, Weibull, and fBM when $Num_{STAs} = 4$.

Based on these results, we concluded that both the traffic pattern among streams in the downlink MU-MIMO channel and the channel conditions affected the system performance.

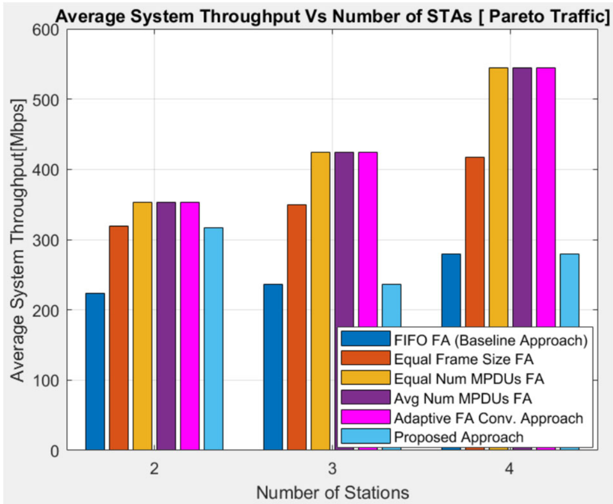
4.5. Performance under Different Number of Stations

The performance of the proposed approach was evaluated under the effects of $Num_{STA} = 4$ and the channel condition SNR = 8 dB for the cases of the Weibull, Pareto, and fBM traffic models. As the results in Figure 5a–c show, when the number of stations ranged from two to four, the system throughput performance was significantly increased in all traffic models, as the traffic rate increased with an increasing number of stations. However, due to the effect of heterogeneous traffic patterns in the different traffic models, the performance of the proposed approach varied even under the same number of stations.

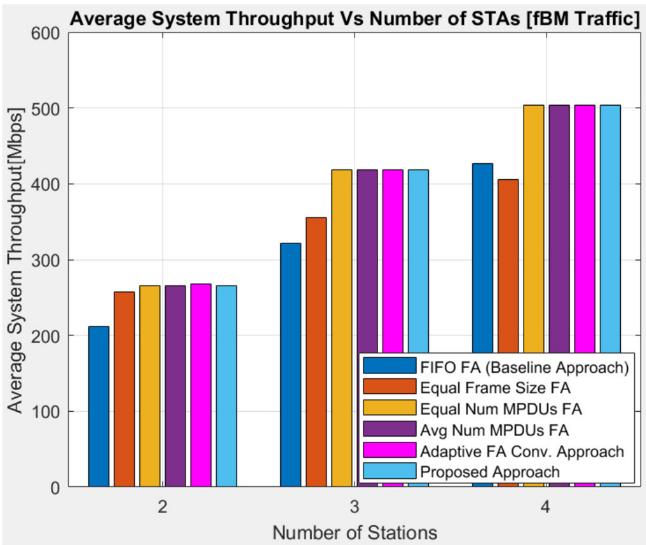
The performance of the proposed adaptive aggregation achieved a better performance than the FIFO FA (baseline approach) due to the adaptive aggregation approach it employed. However, the Adaptive FA Conv. approach achieved a better performance in all scenarios because it only focused on optimizing the system throughput at the expense of increasing the delay. On the contrary, the proposed approach achieved the maximum system throughput performance while considering the lowest cost in terms of increasing the delay that was closest to the maximum performance achieved by the Adaptive FA Conv. approach under the effects of a variable number of STAs.



(**a**) Weibull traffic model



(**b**) Pareto traffic model



(**c**) fBM traffic model

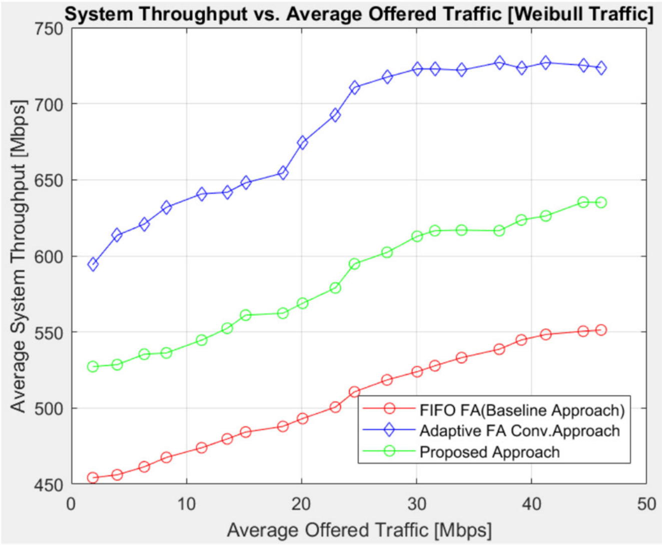
Figure 5. Performance of system throughput versus number of stations when the channel condition was SNR = 8 dB for the Weibull, Pareto, and fBM traffic models.

As the results given in Figure 5 show, the proposed approach always outperformed the FIFO FA (baseline approach), particularly in the Weibull and fBM traffic models, in which the traffic rate variation among streams was less bursty. The proposed approach achieved the maximum performance of 583 Mbps using the Weibull traffic model, whereas a lower performance of 279 Mbps was achieved in the Pareto traffic model with the same number of STAs (four). Likewise, when the number of stations was two, the worst performance of 256 Mbps was achieved by the fBM traffic, but the proposed approach achieved a better performance of 265 Mbps due to its use of an adaptive aggregation approach. These results showed that the number of stations affected the performance of the system throughput behavior under the conditions of heterogeneous traffic patterns among streams in the downlink MU-MIMO channel and under the given channel condition of SNR = 8 dB. In general, according to the results, the performance of the proposed approach was better at coping with less bursty traffic conditions, such as the Weibull and fBM traffic models. This showed that the performance of different aggregation rules adopted in different aggregation policies [10] were affected by the traffic pattern when attempting to provide the optimal minimum delay in the performance of the proposed adaptive aggregation algorithm.

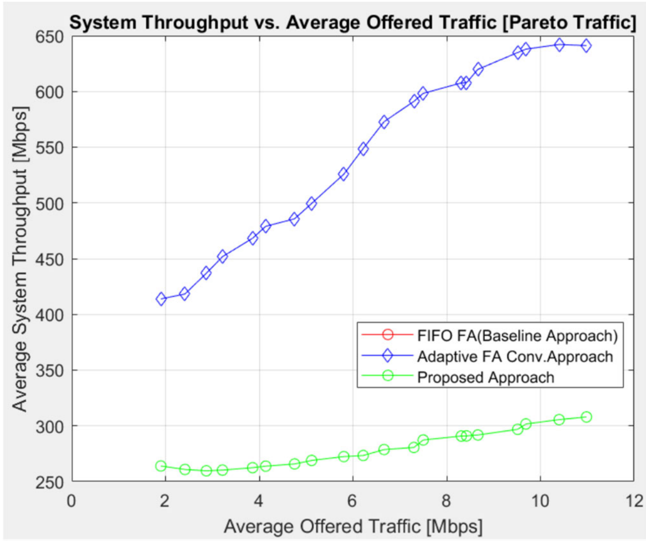
4.6. Performance of System Throughput vs. Average Offered Traffic Load

The results in Figure 6a–c show the performance of system throughput behaviour with increasing average offered traffic load while considering SNR = 8 dB and Num_{STA} = 4 under the effects of different traffic models (Weibull, Pareto, and fBM). This experiment examined the performance of the proposed adaptive aggregation algorithm over the FIFO aggregation algorithm and Adaptive FA Con. approach in terms of the trade-off between maximizing the system throughput and minimizing the system delay.

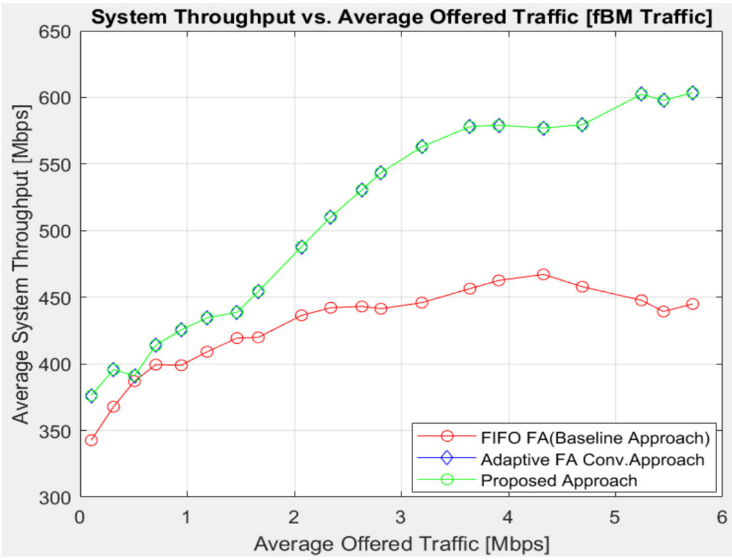
According to the results shown in Figure 6a–c, the proposed approach achieved the maximum performance when the offered frame size was increased in all traffic models. For instance, the proposed approach achieved the maximum performance of 635 Mbps using the Weibull traffic model and 603 Mbps using the fBM traffic model. In the case of the Pareto traffic model, the proposed approach achieved a maximum system throughput performance of 308 Mbps, which is the worst performance as compared to the results achieved by the Weibull and fBM traffic models. This result illustrated that the proposed approach was better at coping with the Weibull and fBM traffic models than the Pareto traffic model. This was because the traffic pattern among streams in the Pareto traffic was highly bursty, thus the FIFO FA (baseline approach) outperformed with a lower cost of queue delay for the optimal system throughput performance. However, the FIFO FA (baseline approach) achieved the worst performance of 550 Mbps in the Weibull traffic model and 308 Mbps in the Pareto traffic model due to the nature of traffic in those models. Therefore, according to these results, our proposed approach was significant and could efficiently maximize the system throughput of the WLAN in the downlink MU-MIMO channel by minimizing the cost of increasing delay. However, according to the results shown in Figure 6a,b, the Adaptive FA Conv. approach always achieved the maximum performance by employing a dynamic adaptive aggregation selection strategy while focusing on optimizing the throughput. However, it provided a suboptimal solution at the expense of maximum delay even though it achieved maximum throughput. In general, this experiment illustrated that the traffic load in different traffic models could affect the performance of a WLAN to realize an optimal trade-off between maximizing throughput and minimizing delay.



(**a**) Weibull traffic model



(**b**) Pareto traffic model



(**c**) fBM traffic model

Figure 6. Performance of system throughput versus average offered traffic load when $Num_{STAs} = 4$ and SNR = 8 dB for the Weibull, Pareto, and fBM traffic models.

5. Conclusions and Future Works

The release of the new IEEE 802.11 standards such as IEEE 802.11ax and IEEE 802.11ay, 5G technologies, and the massive amount of traffic with mixed types of applications generated in modern networks demand different network performance requirements in terms of maintaining some form of an optimal trade-off between maximizing throughput and minimizing delay. However, the majority of existing researchers only attempted to maximize the throughput or minimize the delay. Both the performance of throughput and delay can be affected by several factors such as heterogeneous traffic pattern, target aggregate frame size, channel condition, competing stations, etc. However, under the effects of uncertain conditions of heterogeneous traffic patterns and channel conditions in a network, determining the optimal target aggregate frame size is significant and can be controlled to manage both the throughput and delay. The main contribution of this paper was to propose an adaptive aggregation algorithm aiming to maximize the system throughput performance of a WLAN in the downlink MU-MIMO channel by maintaining the lowest cost in terms of increasing the system delay. Different aggregation algorithms such as FIFO FA, Equal Frame Size FA, Equal MPDUs Agg FA, and Avg Num MPDUs FA were adopted to achieve the dynamic adaptive aggregation selection scheme to realize an optimal trade-off between maximizing throughput and minimizing delay. The effects of channel conditions, heterogeneous traffic patterns for VoIP and video traffic applications, and number of stations were considered when evaluating the proposed approach. Through a simulation, the performance of the proposed approach was evaluated under various channel conditions, traffic patterns with a traffic mix of VoIP and video, and number of STAs. According to the results, the proposed approach achieved a significant performance in terms of both system throughput and system delay over the FIFO FA (baseline approach), and also achieved a better performance closest to the Adaptive FA Conv. approach, which only focuses on achieving the maximum throughput performance with a cost of maximum delay, than the nonadaptive FIFO FA (baseline approach).

Future work will extend the proposed adaptive aggregation approach to operate in real traffic scenarios. We will study the effects of different channel models such as Rayleigh and Rician on both uplink and downlink WLAN channels. Moreover, we will evaluate our approach on an experimental testbed.

Author Contributions: Conceptualization, L.K. and J.D.; Data curation, L.K.; Investigation, L.K.; Methodology, L.K.; Software, L.K.; Supervision, M.D., J.D. and J.C.; Validation, M.D. and J.D.; Writing—original draft, L.K.; Writing—review and editing, M.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are available to any researcher upon request.

Acknowledgments: We would like to thank all authors for their contributions and for the success of this manuscript. Moreover, we would like to thank the editors and anonymous reviewers of this manuscript.

Conflicts of Interest: The authors declare that there is no potential conflict of interest concerning the research, authorship, and/or publication of this article.

References

1. *IEEE Standard 802.11ac*; Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz. IEEE: Piscataway Township, NJ, USA, 2013.
2. Liao, R.; Bellalta, B.; Oliver, M.; Niu, Z. MU-MIMO MAC protocols for wireless local area networks: A survey. *IEEE Commun. Surv. Tutor.* **2014**, *18*, 162–183. [CrossRef]
3. Yin, J.; Wang, X.; Agrawal, D.P. Optimal packet size in error-prone channel for IEEE 802.11 distributed coordination function. In Proceedings of the 2004 IEEE WCNC Wireless Communications and Networking Conference, Atlanta, Georgia, 21–25 March 2004; Volume 3, pp. 1654–1659. [CrossRef]
4. Coronado, E.; Thomas, A.; Riggio, R. Adaptive ML-Based Frame Length Optimization in Enterprise SD-WLANs. *J. Netw. Syst. Manag. Mar.* **2020**, *28*, 850–881. [CrossRef]
5. Deng, J.; Davis, M. An adaptive packet aggregation algorithm for wireless networks. In Proceedings of the 2013 International Conference on Wireless Communications and Signal Processing, Hangzhou, China, 24–26 October 2013; pp. 1–6.
6. Xiao, Y.; Rosdahl, J. Throughput and delay limits of IEEE 802.11. *IEEE Commun. Lett.* **2022**, *6*, 355–357. [CrossRef]
7. Teymoori, P.; Yazdani, N.; Hoseini, S.A.; Effatparvar, M.R. Analyzing delay limits of high-speed wireless ad hoc networks based on IEEE 802.11 n. In Proceedings of the 2010 5th International Symposium on Telecommunications, Tehran, Iran, 4–6 December 2010; pp. 489–495.
8. Zhang, Q.; Iversen, V.B.; Fitzek, F.H. Throughput and delay performance Analysis of Packet Aggregation Scheme for PRMA. In Proceedings of the 2008 IEEE Wireless Communications and Networking Conference (WCNC 2008), Las Vegas, NV, USA, 31 March–3 April 2008; pp. 1380–1384.

9. Moriyama, T.; Yamamoto, R.; Ohzahata, S.; Kato, T. Frame aggregation size determination for IEEE 80211ac WLAN considering channel utilization and transfer delay. In Proceedings of the WINSYS 2017—14th International Conference on Wireless Networks and Mobile System, Madrid, Spain, 26–28 July 2017; Volume 6, pp. 89–94. [CrossRef]
10. Kassa, L.; Davis, M.; Cai, J.; Deng, J. A new adaptive frame aggregation method for downlink WLAN MU-MIMO channels. *J. Commun.* **2021**, *16*, 311–322. [CrossRef]
11. Kassa, L.; Davis, M.; Deng, J.; Cai, J. Performance of an Adaptive Aggregation Mechanism in a Noisy WLAN Downlink MU-MIMO Channel. *Electronics* **2022**, *11*, 754. [CrossRef]
12. Ci, S.; Sharif, H. Adaptive optimal frame length predictor for IEEE 802.11 wireless LAN. In Proceedings of the 6th International Symposium on Digital Signal Processing for Communication Systems (DSPCS'2002), Sydney, Australia, 28–31 January 2002; pp. 59–63.
13. Bianchi, G. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE J. Sel. Areas Commun.* **2000**, *18*, 535–547. [CrossRef]
14. Lin, P.; Lin, T. Machine-learning-based adaptive approach for frame-size optimization in wireless LAN environments. *IEEE Trans. Veh. Technol.* **2009**, *58*, 5060–5073. [CrossRef]
15. Modiano, E. An adaptive algorithm for optimizing the packet size used in wireless ARQ protocols. *Wirel. Netw.* **1999**, *5*, 279–286. [CrossRef]
16. Kim, S.; Yun, J.H. Efficient Frame Construction for Multi-User Transmission in IEEE 802.11 WLANs. *IEEE Trans. Veh. Technol.* **2019**, *68*, 5859–5870. [CrossRef]
17. Bellalta, B.; Barcelo, J.; Staehle, D.; Vinel, A.; Oliver, M. On the performance of packet aggregation in IEEE 802.11ac MU-MIMO WLANs. *IEEE Commun. Lett.* **2012**, *16*, 1588–1591. [CrossRef]
18. Lin, C.H.; Chen, Y.T.; Lin, K.C.J.; Chen, W.T. acPad: Enhancing channel utilization for 802.11 ac using packet padding. In Proceedings of the IEEE INFOCOM 2017-IEEE Conference on Computer Communications, Atlanta, GA, USA, 1–4 May 2017; pp. 1–9. [CrossRef]
19. Lin, C.H.; Chen, Y.T.; Lin, K.C.J.; Chen, W.T. Fdof: Enhancing channel utilization for 802.11 ac. *IEEE/ACM Trans. Netw.* **2018**, *26*, 465–477. [CrossRef]
20. Nomura, Y.; Mori, K.; Kobayashi, H. High-efficient frame aggregation with frame size adaptation for downlink MU-MIMO wireless LANs. *IEICE Trans. Commun.* **2016**, *E99B*, 1584–1592. [CrossRef]