Efficient Frame Construction for Multi-User Transmission in IEEE 802.11 WLANs

Sanghyun Kim, and Ji-Hoon Yun, *Member, IEEE*

***Abstract*—The latest standard of IEEE 802.11 WLANs embraces multi-user (MU) transmission via OFDMA, MU-MIMO or a mixture of both. However, due to the frame-basis transmission of WLANs, all concurrent user frames must have an identical transmission duration so as to fit in a common MU frame, which requires shorter frames padded with dummy bits, thus reducing transmission efficiency. Since stations are likely to have heterogeneous traffic demands and transmission bit rates in real networks, such inefficiency arises as a practical problem. To address this challenge, first we analyze how different overhead components impact the efficiency of MU transmission. Then, we propose a scheme to construct an MU frame with an optimal length maximizing its transmission efficiency, which provides a unified framework for OFDMA and MU-MIMO in both downlink and uplink. Through simulation, we demonstrate that the proposed scheme integrated with various scheduling algorithms reduces transmission delay and enhances traffic delivery ratio considerably compared to basic approaches.**

***Index Terms*—WLANs, IEEE 802.11, multi-user transmission, OFDMA, MU-MIMO, frame construction**

I. Introduction

Multi-user (MU) transmission technology such as orthogonal frequency-division multiple access (OFDMA) and multi-user multiple-input and multiple-output (MU-MIMO) enables concurrent transmission of multiple user frames from a single transmitter to a group of receivers (for downlink) and vice versa (for uplink) to achieve higher transmission efficiency and statistical multiplexing gains. Owing to such advantages, MU transmission has been adopted by many *de facto* wireless communication systems. In particular, the latest wireless LAN (WLAN) standard *IEEE 802.11ax* [1] supports MU transmission based on OFDMA, MU-MIMO or a mixture of both. The access point (AP) of IEEE 802.11ax can transmit up to 9 subcarrier groups in a 20MHz channel and 8 concurrent spatial streams allocated to individual stations, thus achieving a high degree of multi-user multiplexing.

However, the advantages of MU transmission is not always achievable in WLANs due to frame-basis transmission.[1] In WLANs, an MU frame is constructed such that all multiplexed user frames have an identical transmission duration, thus fitting in the common frame. Therefore, if the airtime required to transmit a user frame is shorter than those of the others, the frame pads dummy bits, called *padding bits*, at the tail until it has the same transmission duration as the others, thus degrading transmission efficiency. As we show through simulation in the performance evaluation section (Section VII), the impact of the padding overhead of an MU frame on performance is considerable. We note that the problem is common and serious in practical network environments that stations have heterogeneous traffic demands and transmission bit rates in general. Therefore, it should be considered as an important design factor of MU transmission in WLANs.

Despite the importance of MU frame construction, the problem has not been explored much in the literature. Scheduler designs for MU transmission focus on user selection only [2]–[8], align all multiplexed user frames to the shortest one [9] [10] or to the average [11] (which we classify into *Min-aligned* and *Avg-aligned* schemes, respectively), or allow as much data as possible to be conveyed by a user frame, i.e., align to the longest user frame [12] (which we call a *Max-aligned* scheme). Our evaluation shows that such simple methods fail to perform best in all network conditions because the problem is not straightforward; shortening the length of an MU frame will reduce padding bits, but at the expense of less data to be conveyed and a higher ratio of other overhead (preamble, header, etc.). There has also been an approach to pad other stations' data instead of dummy bits [13] [14], which, however, requires standard modification and increases the complexity of both transmission and reception processes.

In this paper, we propose a new data frame construction scheme for MU transmission in IEEE 802.11 WLANs. The proposed scheme finds the length of an MU frame to maximize its transmission efficiency (the amount of data conveyed by a frame in unit time), taking into account the status of buffers and transmission bit rates of stations. To this end, first we analyze how different overhead components impact the efficiency of MU transmission. Then, we formulate the MU frame construction problem as an optimization problem and show its $\mathcal{NP}$-completeness. The proposed scheme is designed

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (2017R1D1A1A09000986) and the Technology Development Program (C0503976) funded by the Ministry of SMEs and Startups (MSS, Korea).

Corresponding author: Ji-Hoon Yun.

Sanghyun Kim is with the Department of Electrical and Information Engineering and the Research Center for Electrical and Information Technology, Seoul National University of Science and Technology, Seoul, Korea (e-mail: shk0787@seoultech.ac.kr).

Ji-Hoon Yun is with the Department of Electrical and Information Engineering and the Research Center for Electrical and Information Technology, Seoul National University of Science and Technology, Seoul, Korea (e-mail: jhyun@seoultech.ac.kr).

[1]This issue is unique in MU transmission of WLANs due to a varying frame length.

to find a suboptimal solution of the problem via decomposition and construct an MU frame accordingly as a unified framework of MU frame construction for OFDMA and MU-MIMO in both downlink (DL) and uplink (UL) transmission directions. The scheme is combined with a general scheduling algorithm that allocates radio resources to a set of stations for each frame transmission based on the algorithm-specific objective; the decision is then passed to the scheme and the final frame is constructed.

Through simulation, we evaluate the proposed scheme in comparison with basic schemes—Min, Avg and Max-aligned—when integrated with three scheduling algorithms (round-robin, proportional fair and queue length-weighted scheduling) and demonstrate that it reduces transmission delay by up to 78% (single BSS case) and improves traffic delivery ratio by up to 107% (overlapping BSSs case).

The rest of this paper is organized as follows. Section II reviews related work and Section III explains MU transmission in IEEE 802.11ax WLANs. Section IV illustrates how data frame construction affects transmission efficiency in two-station case. The system model under consideration is given in Section V and the details of the proposed scheme is described in Section VI. Section VII shows simulation results for performance evaluation and Section VIII concludes the paper.

II. Related Work

Unlike the conventional single-user transmission, the performance of MU transmission highly depends on how users and radio resources are scheduled. Thus a rich body of work has tackled the scheduling problem, mostly for MU-MIMO. User selection was developed in various aspects; via distributed feedback contention [3], with no use of channel state information to avoid bandwidth overhead [15] [16], for large-scale DL MU-MIMO [5], considering interference between users [17] and interference of neighbouring APs additionally [8], and jointly determining the time interval of channel sounding [7]. AP-station association was also considered together, taking channel correlation between stations into account [2] and via an auction-based approach [18]. Scheduler designs were proposed for fairness with minimum loss of throughput [4]; implementing a scheduler in an approximating greedy algorithm was also proposed to have low complexity [19]. A few research considered OFDMA; Bankov *et al.* [6] claimed that existing schedulers are inefficient for UL OFDMA and proposed a new one. The above proposals, however, focused on user set selection and resource allocation only, and did not consider the details of frame construction.

Some work on scheduling considered MU frame construction. Nomura *et al.* [9] proposed to limit the size of user frames to be multiplexed to enhance throughput and lower frame errors in DL MU-MIMO, thus corresponding to Min-aligned. A simple scheduling scheme combined with Min-aligned prioritizing the stations with longer backlogged queues was proposed in [10]. Another proposal was to align the lengths of user frames to their average [11], which is Avg-aligned. Max-aligned was also considered in scheduling with consideration of correlation metrics between stations [12].

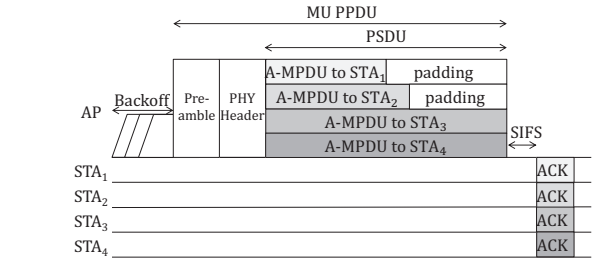
There has been a notable research work focusing on the padding problem of DL MU-MIMO. Lin *et al.* proposed to replace padding bits of a station's frame with data of other stations, thus increasing the transmission efficiency of an MU frame [13] [14]. However, this approach requires a considerable modification of the standards to allow multiple destinations within a spatial stream or equivalently changing a set of multiplexed user frames and accompanied modulation and coding schemes in the middle of an MU frame. Such a modification also increases complexity in both transmission and reception processes.

There have been some research works on multi-user transmission in IEEE 802.11ax WLANs. A simulation platform to evaluate the features including multi-user transmission (OFDMA and MU-MIMO) along with non-continuous channel bonding and link adaptation was designed and implemented [20]. Analytical models of random access for UL OFDMA were developed in [21] and [22] to find the impact of various configuration parameters on system efficiency. Hoefel *et al.* [23] proposed to use uplink power control to mitigate the adverse effects of in-phase and quadrature imbalances in UL MU-MIMO of 802.11ax.

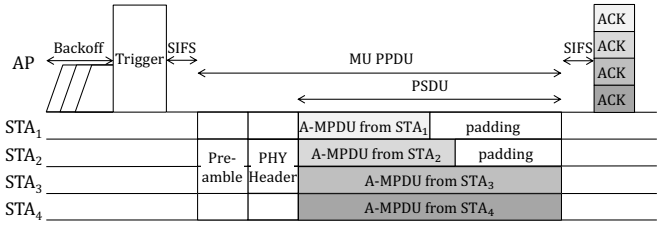
Some studies focused on the design of MAC-layer procedures for MU-MIMO. The key requirements of a MAC design were provided in [24]. Various MAC designs were proposed for a new backoff procedure [25], for a fair MAC protocol [26], and to solve hidden terminal problems using precoding vectors from zeroforcing [27]. A distributed MAC protocol that achieves both multiplexing and receive diversity gains was proposed in [28]. An AP-controlled MAC with differentiation of traffic classes was developed [29]. TXOP sharing was also enhanced [30].

There have also been many studies to evaluate and model the performance of MU-MIMO in WLANs. Comparisons of various transmission methods including SISO and MU-MIMO [31] and different precoding types [32] were made. Redieteab *et al.* [33] modelled both PHY and MAC operation of MU-MIMO for evaluation, thus obtaining more accurate performance values. Testbed experiments of MU-MIMO with commercial APs found that non-MU-MIMO stations suffer in an MU-MIMO-activated WLAN [34]. Thanh *et al.* [35] proposed a modified MAC layer control frame including those for channel reservation and analyzed the performance using Markov chains. Ma *et al.* [36] showed that there is an optimal sounding interval for single-user transmit beamforming and MU-MIMO. A closed-form expression of rate reduction due to channel quantization error was derived [37].

There have been an increasing number of research works on multi-user transmission in full-duplex (FD) WLANs. In FD WLANs, a medium access control (MAC) protocol plays an important role to maximize FD transmission opportunities by finding a pair of STAs to be multiplexed within a FD transmission. Many designs exploited the RTS/CTS exchange mechanism to find such a pair having similar traffic demands [38]–[41]. In order to find a STA pair with no inter-STA interference, many designs let an AP to construct a map of inter-STA interference from reported or overheard information [39], [42]. In [43], distributed power control was considered



(a) Downlink



(b) Uplink

Fig. 1. Frame format and exchange procedure of multi-user transmission in IEEE 802.11ax WLANs

to manage inter-STA interference. Chen *et al.* [44] proposed to use probabilistic access differentiation based on the level of inter-STA interference. Multiplexing more STAs than two within a FD transmission has also been studied. In [45], each STA is assigned a subcarrier index and informed of a UL grant based on it during an arbitration period, thereby avoiding collision and maximizing FD opportunities in dense AP environments. Lee *et al.* [46] proposed in-frame querying to best utilize the airtime of a FD transmission by allowing another STA to start a new UL transfer in the middle of a FD frame once the transfer of an assigned STA finishes before the end of the frame.

III. Multi-User Transmission and Its Overhead in IEEE 802.11ax WLANs

For MU transmission in IEEE 802.11ax, an MU frame is constructed to multiplex up to nine *resource units* (RUs), each of which is a distinct group of contiguous subcarriers, in a 20MHz channel or eight spatial streams (SSs) destined to individual stations (STAs). Fig. 1 illustrates the MU transmission procedure of IEEE 802.11ax for four stations (STAs), which is applied to both OFDMA and MU-MIMO transmission. In the downlink (DL) case, upon the end of a backoff procedure, a frame transmission starts with a preamble in the beginning followed by a physical-layer (PHY) header. Then, the physical layer convergence protocol service data unit (PSDU), which is the aggregate MAC protocol data unit (A-MPDU)[2] with padding bits, is mapped to RU or SS by resource allocation of an AP's scheduler and transmitted at STA-specific transmission bit rate. We call such a granularity of resource allocation for MU transmission *resource granularity* (RG) in common (corresponding to either RU or SS) in the paper. Then, individual acknowledgement (ACK) transmissions of receiving STAs follow in short interframe space (SIFS). In the uplink (UL) case, a trigger frame is transmitted first by AP at the end of a backoff procedure and the rest of the procedure, which is almost same as DL's, follows.

In the MU transmission, we consider any component except the transmission part of A-MPDU as the overhead that reduces physical-layer transmission efficiency. We classify the overhead into two types: *protocol overhead* and *padding bits*. The protocol overhead results from the information needed to exchange with the counterpart(s) of communication as defined in the standard. As shown in Fig. 1, the protocol overhead is the set of components except PSDU, which are transmitted at the lowest transmission bit rate or one of the basic rate set (BRS) and thus their airtime consumption is mostly fixed. We denote the airtime consumption of the protocol overhead by T_o, which is given below:

$$T_o = \begin{cases} T_b + T_{pre} + T_h + T_{sifs} + T_{ack} & \text{DL} \\ T_b + T_{tr} + T_{sifs} + T_{pre} + T_h + T_{sifs} + T_{ack} & \text{UL} \end{cases} \quad (1)$$

where T_b, T_{pre}, T_h, T_{sifs}, T_{ack} and T_{tr} are the backoff, airtimes of preamble, PHY header, SIFS, ACK and trigger frames, respectively.

As shown in the figure, if there exist different amounts of data to transmit to or from individual STAs and heterogeneous transmission bit rates are used for them, which results in different data transmission durations between them in the frame, then gaps between these transmission durations are filled with padding bits such that each PSDU (A-MPDU + padding bits) occupies the same number of symbols, but at the expense of the waste of airtime. As the amount of such padding bits gets larger, the efficiency of MU transmission decreases. In the real world, STAs have heterogeneous traffic demands and transmission bit rates. Therefore, the inefficiency due to padding bits is somewhat inevitable.

In order to minimize the impact of the protocol overhead on transmission efficiency, we have to increase the length of PSDU as long as possible so that the maximum amount of data can be conveyed by the frame. To the contrary, to minimize padding bits, we have to decrease the length of PSDU. That is, there exists a trade-off in determining the length of PSDU which impacts transmission efficiency and we focus on this trade-off problem in the paper.

IV. Frame Construction and Transmission Efficiency: Illustration of Two-Station Case

In this section, we illustrate how data frame construction impacts the efficiency of MU transmission in more detail via the analysis of a simple two-station case.

Let l_i be the length of the A-MPDU addressed to or transmitted by STA i in an MU data frame. Denote the transmission bit rate for STA i by r_i. Then, the transmission duration of STA i's A-MPDU is given as l_i/r_i.

In the two-station case (with STAs 1 and 2), we assume that the A-MPDU transmission duration of STA 2 is not shorter than that of STA 1, i.e., $l_1/r_1 \leq l_2/r_2$, thus the A-MPDU of STA 1 is padded with dummy bits such that its airtime is aligned with l_2/r_2. Then, the throughput of the transmission,

[2]A-MPDU is the concatenation of multiple MPDUs.

which we denote by G, is obtained as

$$G = \frac{l_1 + l_2}{T_o + l_2/r_2}. \tag{2}$$

We rewrite it as

$$G = \frac{l_1 + l_2}{l_2 + r_2 T_o} r_2 = r_2 + \frac{l_1 - r_2 T_o}{l_2 + r_2 T_o} r_2 \tag{3}$$

where $r_2 T_o$ can be interpreted as the amount of the protocol overhead in terms of bits.

Throughput G shows different trends for l_2 depending on the relationship between l_1 and $r_2 T_o$, which is classified into the following two cases:

1) $l_1 \leq r_2 T_o$: The second term of Eq. (3) is negative and thus throughput is a monotonically increasing function of l_2, which implies that l_2 can be set to the maximum data size available. The throughput, however, is always lower than the transmission bit rate of STA 2. In this case, the waste of radio resource consumed for the protocol overhead is larger than the overhead of padding bits. So, we better transmit as much data as possible at once.
2) $l_1 > r_2 T_o$: The second term is now positive and throughput is a monotonically decreasing function of l_2, thus maximized when l_2 gets as small as possible. So, the maximum throughput is achieved when l_2 is set to $l_1 r_2 / r_1$ (since $l_1/r_1 \leq l_2/r_2$) and given as $\frac{l_1(r_1+r_2)}{l_1+r_1 T_o}$. In this case, l_1 is large enough to make up for the protocol overhead. So, we better focus on limiting padding bits.

The backoff time is included in T_o which appears in the denominator of the throughput calculation in Eq. (2). Thus, if the backoff time gets longer, T_o increases and the resulting throughput decreases. According to the above analysis, increasing T_o expands the range of l_1 meeting the above case 1) where conveying more data in a frame is better, thus the impact of T_o on throughput is reduced.

In what follows, we consider a general case with many STAs. In such a case, determination of a frame length is not as simple as the two-station case and thus we focus on finding a general rule.

V. System Model

We consider a WLAN network composed of APs and multiple STAs; each AP serves STAs connected to it, forming a basic service set (BSS). The MAC and PHY behavior including MU transmission under consideration is based on IEEE 802.11ax. Traffic is generated and buffered at per-STA queues of AP (DL) or the queue of a STA's own (UL). The buffer status of each STA is known to its serving AP via buffer status reports (BSRs) to assist the AP in allocating UL MU resources.[3]

An MU transmission requires the following two steps to be done before transmission: (1) resource allocation; and (2) data frame construction. The scheduler of the AP selects a set of STAs for their user frames to be multiplexed in an MU frame and allocates RGs to them based on a pre-defined objective. Then, the determined STA set and resource allocation result is passed to the data frame construction process. We assume that the basic scheme as a baseline constructs an MU frame such that the transmission duration of each STA's frame is aligned to the maximum (Max-aligned), average (Avg-aligned) or minimum (Min-aligned) among the multiplexed frames.

In order to make RG allocation, a scheduler determines a STA to use each RG under the constraint that the final set of users meets the orthogonality between them (no inter-RG interference), i.e., a given transmission bit rate for a STA results in no transmission failure unless a collision occurs. Once allocation of a RG to a STA is made, the status information of all STAs (e.g. past throughput for proportional fair scheduling) is updated and the scheduler goes through the same procedure for the next RG until all RGs are allocated. The transmission bit rate of a STA is readily determined based on IEEE 802.11ax's MCS table once resource allocation is finished (the MCS of each STA for a single RG is given *a priori* according to a randomly-assigned link quality).

For determination of a STA to be allocated each RG, we consider three types of scheduling algorithms: round robin (RR), proportional fair (PF) and queue length-weighted (QL). With RR, the scheduler allocates STAs sequentially. With PF, the scheduler sorts and picks STAs in the order of $s_i/\bar{s}_i$ where s_i is the expected throughput of STA i through this transmission and $\bar{s}_i$ is the average throughput achieved by STA i with a weight on recent samples. With QL, the scheduler sorts and picks STAs in the order of $q_i/\bar{s}_i$ where q_i is the amount of data queued for STA i; $q_i/\bar{s}_i$ corresponds to the total airtime needed to flush all the queued data for the STA. Thus, QL tries to minimize the maximum delay experienced by STAs.

VI. Data Frame Size Control

In this section, we formulate the data frame construction problem and solve it. We then develop an algorithm that finds a solution efficiently.

A. Problem Formulation

Let T_p be the transmission duration of PSDU. When N user frames are multiplexed in an MU frame transmission, the achieved throughput is given as

$$G = \frac{\sum_{i=1}^{N} l_i}{T_o + T_p} \tag{4}$$

where l_i is the length (data bits) of A-MPDU for STA i. If A-MPDU conveying all queued data for STA i (amounting to q_i) can be transmitted within T_p, A-MPDU is constructed as such. If not, A-MPDU is constructed so as to convey as much data as possible while having its transmission time as T_p. Without loss of generality, we index STAs in the ascending order of q_i/r_i, i.e., $q_i/r_i \leq q_{i+1}/r_{i+1}$. For simplicity of exposition, we refer to the amount of data stored in a queue, denoted as q_i for STA i, as the total number of data bits.

[3]In IEEE 802.11ax, a STA reports its buffer status to AP using either the QoS Control field or the BSR Control subfield of frames it transmits.

Then, the target problem, denoted by P, is expressed as

$$\begin{aligned} P: \quad & \max_{\{l_i\},T_p} G \\ & \text{s.t.} \\ & l_i/r_i \le T_p, \\ & l_i \le q_i, \\ & i = 1, \ldots, N. \end{aligned} \tag{5}$$

Proposition 1. *P is $\mathcal{NP}$-complete.*

Proof. According to Freund and Jarre's work [47], the *sum-of-ratios problem* is $\mathcal{NP}$-complete, whose general form is expressed as

$$\begin{aligned} & \max_{\mathbf{x}} h(\mathbf{x}) + \sum_{j=1}^{p} \frac{f_j(\mathbf{x})}{g_j(\mathbf{x})} \\ & \text{s.t.} \\ & \mathbf{x} \in \mathcal{S} \subset \mathbb{R}^n \end{aligned} \tag{6}$$

when $f_j(\mathbf{x}) \ge 0$, $g_j(\mathbf{x}) > 0$, h and f_j are concave and g_j is convex for all j and $\mathcal{S}$ is a convex set; $\mathbf{x}$ is a vector. P is the sum-of-ratios problem with $\mathbf{x} = [l_1, \ldots, l_N, T_p]$, $f_j(\mathbf{x}) = l_j \ge 0$, $g_j(\mathbf{x}) = T_o + T_p > 0$ and the constraints form a convex set. □

The $\mathcal{NP}$-completeness of P suggests that there is no efficient polynomial-time algorithm to solve it exactly. Thus it is prohibitive to find the global optimizer in terms of computational complexity. Our approach to find a suboptimal solution of P with low computational complexity is two-fold: (1) transform P into an equivalent problem with reduced problem variables; and then (2) decompose the new problem into subproblems and solve them sequentially to find a suboptimal solution.

The solution of P is given as $(l_1, l_2, \cdots, l_N, T_p)$. Assume that STA j^*'s transmission time is aligned with T_p, i.e., $T_p = l_{j^*}/r_{j^*}$. Then, for STA $i < j^*$, we have $l_i = q_i$ and, for $i > j^*$, $l_i = r_i T_p$.[4] This means that P is equivalent to finding a two-tuple of (j^*, l_{j^*}). Then, P is rewritten in its equivalent form denoted by P' as

$$\begin{aligned} P': \quad & (j^*, l_{j^*}) = \arg\max_{j, l_j} G \\ & \text{s.t.} \\ & r_j q_{j-1}/r_{j-1} \le l_j \le q_j, \\ & j = 1, \ldots, N \end{aligned} \tag{7}$$

whose solution space is $N \times \mathbb{R}$ (reduced from $\mathbb{R}^{N+1}$ of P).

In order to find a suboptimal solution of P', we first (1) find STA j such that T_p is aligned with q_j/r_j and then (2) adjust l_j for further optimization. We divide P' as such into two subproblems:

$$\begin{aligned} P1: \quad & j^* = \arg\max_{j} G(j) \\ & \text{s.t.} \\ & l_i = q_i, && \text{for } i \le j, \\ & l_i/r_i = l_j/r_j = T_p, && \text{for } i > j, \\ & i = 1, \ldots, N, j = 1, \ldots, N \end{aligned} \tag{8}$$

[4]Then, the assumption $q_i/r_i \le q_{i+1}/r_{i+1}$ leads to $l_i/r_i \le l_{i+1}/r_{i+1}$.

and

$$\begin{aligned} P2: \quad & l_{j^*} = \arg\max_{l_{j^*}} G(j^*) \\ & \text{s.t.} \\ & q_{j^*-1}/r_{j^*-1} \le l_{j^*}/r_{j^*} \le q_{j^*}/r_{j^*}. \end{aligned} \tag{9}$$

B. Problem Solution

Assume that the transmission durations of all user frames (PSDUs) are aligned with the one for STA j, i.e., l_j/r_j. Thus, the A-MPDU of STA $i < j$ is padded with dummy bits amounting to $l_j/r_j - l_i/r_i$ and that of STA $i > j$ is limited by l_j/r_j, i.e., $l_i/r_i = l_j/r_j$ which leads to $l_i = r_i(l_j/r_j)$. Then, we rewrite throughput G of Eq. (4) as a function of j as below:

$$\begin{aligned} G(j) &= \frac{\sum_{i=1}^{j-1} l_i + (l_j/r_j)\sum_{i=j}^{N} r_i}{T_o + l_j/r_j} \\ &= \frac{\sum_{i=1}^{j-1} l_i + (T_o - T_o + l_j/r_j)\sum_{i=j}^{N} r_i}{T_o + l_j/r_j} \\ &= \sum_{i=j}^{N} r_i + \frac{\sum_{i=1}^{j-1} l_i - T_o \sum_{i=j}^{N} r_i}{T_o + l_j/r_j}. \end{aligned} \tag{10}$$

In what follows, we observe how $G(j)$ changes as j increases. In order to do that, we first obtain $G(j+1) - G(j)$ as

$$\begin{aligned} & G(j+1) - G(j) \\ &= \sum_{i=j+1}^{N} r_i + \frac{\sum_{i=1}^{j} l_i - T_o \sum_{i=j+1}^{N} r_i}{T_o + l_{j+1}/r_{j+1}} \\ &\quad - \left(\sum_{i=j}^{N} r_i + \frac{\sum_{i=1}^{j-1} l_i - T_o \sum_{i=j}^{N} r_i}{T_o + l_j/r_j} \right) \\ &= \frac{\sum_{i=1}^{j} l_i - T_o \sum_{i=j+1}^{N} r_i}{T_o + l_{j+1}/r_{j+1}} \\ &\quad - r_j - \frac{\sum_{i=1}^{j-1} l_i - T_o \sum_{i=j}^{N} r_i}{T_o + l_j/r_j} \\ &= \frac{\sum_{i=1}^{j} l_i - T_o \sum_{i=j+1}^{N} r_i}{T_o + l_{j+1}/r_{j+1}} \\ &\quad - \frac{r_j(T_o + l_j/r_j) + \sum_{i=1}^{j-1} l_i - T_o \sum_{i=j}^{N} r_i}{T_o + l_j/r_j} \\ &= \frac{\sum_{i=1}^{j} l_i - T_o \sum_{i=j+1}^{N} r_i}{T_o + l_{j+1}/r_{j+1}} - \frac{\sum_{i=1}^{j} l_i - T_o \sum_{i=j+1}^{N} r_i}{T_o + l_j/r_j} \end{aligned} \tag{11}$$

where the numerators of the final two terms are identical. We denote this common numerator by $F(j)$, i.e.,

$$F(j) := \sum_{i=1}^{j} l_i - T_o \sum_{i=j+1}^{N} r_i \tag{12}$$

and finally obtain $G(j+1) - G(j)$ as

$$G(j+1) - G(j) = \frac{F(j)(l_j/r_j - l_{j+1}/r_{j+1})}{(T_o + l_{j+1}/r_{j+1})(T_o + l_j/r_j)}. \tag{13}$$

Proposition 2. *F is a monotonically increasing function of j.*

Proof. In Eq. (12), as j increases, the first term increases and the second term decreases, both monotonically. In other words, $F(j)$ is the sum of two monotonically increasing functions of

j. Therefore, $F(j)$ itself is also a monotonically increasing function of j. □

Proposition 3. *G is a monotonically increasing function of j when $F \leq 0$ and a monotonically decreasing function of j when $F \geq 0$, i.e,*

$$G(j+1) - G(j) = \begin{cases} \geq 0 & F(j) \leq 0 \\ \leq 0 & F(j) \geq 0 \end{cases} \quad (14)$$

for $1 \leq j \leq N-1$.

Proof. In Eq. (13), the denominator is positive and the numerator excluding $F(j)$ is non-positive due to the assumption $l_j/r_j \leq l_{j+1}/r_{j+1}$, which leads to Eq. (14). □

Recall that our first problem $(P1)$ is to find j which maximizes $G(j)$. From the above propositions, we now obtain a proposition regarding the optimum of G as follows.

Proposition 4. *Let j^* be defined as*

$$j^* = \arg \max_{1 \leq j \leq N} G(j). \quad (15)$$

Then, j^* is the smallest j that makes $F(j)$ non-negative. If there is no such j, $j^* = N$.

Proof. We divide the values of F into three cases: (1) $F(1) \geq 0$; (2) $F(1) \leq 0$ and $F(N) \leq 0$; or (3) $F(1) \leq 0$ and $F(N) \geq 0$, and find the optimum of G for each case in the following.

(1): $F(j) \geq 0$ for all j since F is a monotonically increasing function of j from Proposition 2. Then, from Proposition 3, $G(j+1) - G(j) \leq 0$ for all j and thus $G(1)$ is the maximum.

(2): $G(j+1) - G(j) \geq 0$ for all j and $G(N)$ is the maximum. If $F(1) \leq 0$ and $F(N) \leq 0$, $G(j+1) - G(j) \geq 0$ for all j and $G(N)$ is the maximum.

(3): There exists j^* which meets $F(j^*) \leq 0$ and $F(j^*+1) \geq 0$. Then, $G(j^*+1) - G(j^*) \geq 0$ and $G(j^*+2) - G(j^*+1) \leq 0$; thus $G(j^*+1)$ is the maximum.

The above observations prove the proposition. □

Once j^* is found, we can further adjust l_{j^*} while meeting $l_{j^*-1}/r_{j^*-1} \leq l_{j^*}/r_{j^*} \leq q_{j^*}/r_{j^*}$. If $l_{j^*}/r_{j^*} < q_{j^*}/r_{j^*}$, STA j^*'s frame does not contain the whole amount of backlogged data q_{j^*} Therefore, there still remains the problem to find optimal l_{j^*} in $[l_{j^*-1} r_{j^*}/r_{j^*-1}, q_{j^*}]$. The solution of the problem is found as follows.

Proposition 5. *When j^* is given, $G(j^*)$ is maximized when $l_{j^*} = q_{j^*}$.*

Proof. From Eqs. (10) and (12), we have $G(j^*)$ as

$$G(j^*) = \sum_{i=j^*}^{N} r_i + \frac{F(j^*-1)}{T_o + l_{j^*}/r_{j^*}} \quad (16)$$

where the second term of the right side is negative since j^* is the smallest integer value to make $F(j^*)$ non-negative. Therefore, $G(j^*)$ is maximized when this negative term is minimized, which is achieved when l_{j^*} is chosen as the largest possible value, i.e., q_{j^*}. □

Now we can understand the problem solution intuitively. Similarly with the illustration of the two-station case in Section IV, we can consider two cases depending on the value of $F(j)$; l_1 and $r_2 T_o$ of the two-station case are now mapped to $\sum_{i=1}^{j} l_i$ and $T_o \sum_{i=j+1}^{N} r_j$, respectively (for $N = 2$, the general case terms match the two-station case exactly). Therefore, we can assume that $T_o \sum_{i=j+1}^{N} r_j$ corresponds to the protocol overhead. Then, based on the understanding of the two-station case, when $F(j) \leq 0$ (corresponding to Case 1) of the two-station case), the waste of radio resource due to the protocol overhead is dominant and we better increase the amount of data to transmit at the expense of increasing padding bits, which is aligned with the finding for the general case (Proposition 3). In the opposite case ($F(j) > 0$), the overhead of padding bits is dominant and reducing the MU frame length is better (corresponding to Case 2) of the two-station case).

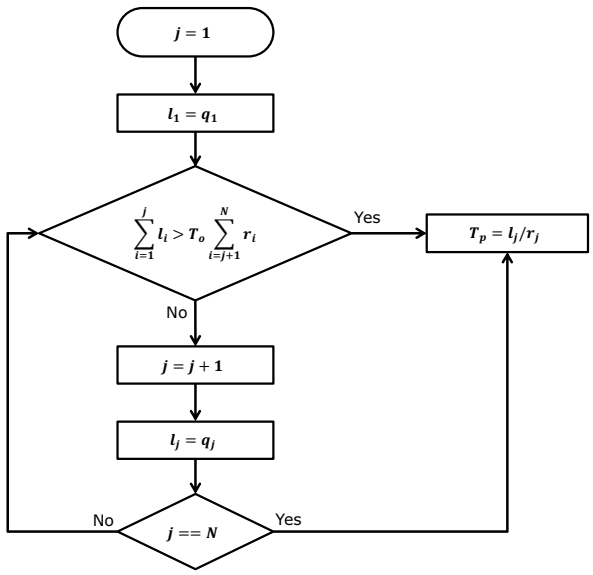


Fig. 2. DFSC algorithm

C. Algorithm Design

From the findings obtained above, we design an algorithm to determine the best data frame length for a given set of selected STAs and resource allocation, which we call the **D***ata* **F***rame* **S***ize* **C***ontrol scheme* (DFSC). The algorithm of DFSC is depicted in Fig. 2. First, STAs are sorted in the ascending order of q_i/r_i, i.e., $q_i/r_i \leq q_{i+1}/r_{i+1}$ for all i where q_i is the queue length of STA i and r_i is the transmission bit rate. Then, the algorithm starts with STA 1 ($j = 1$) setting $l_1 = q_1$ and checks if the STA satisfies the condition $\sum_{i=1}^{j} l_i > T_o \sum_{i=j+1}^{N} r_i$ which has been defined as $F(j)$ in Eq. (12). If not, the algorithm repeats the above step for the next STA. That is, the algorithm finds the first STA making $F(j)$ non-negative. Then, the first j satisfying the condition, also denoted by j^*, is the one maximizing throughput according to Proposition 4. Finally, the frame length is set to q_{j^*}/r_{j^*} according to Proposition 5. If the condition is not met for the last STA (STA N), the algorithm sets the frame length as q_N/r_N since throughput $G(j)$ is still a monotonically increasing function and maximal at N

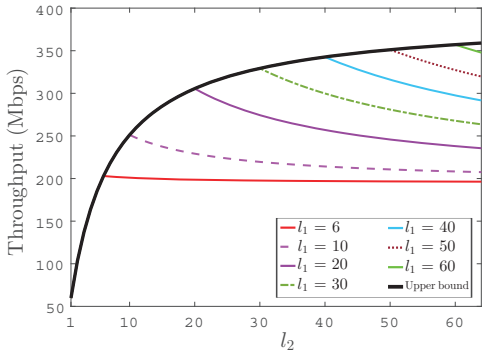


Fig. 3. Throughput for various combinations of A-MPDU sizes in two-STA case (l_1 and l_2 for STAs 1 and 2, respectively, in terms of the number of MPDUs, $l_2 \geq l_1$)

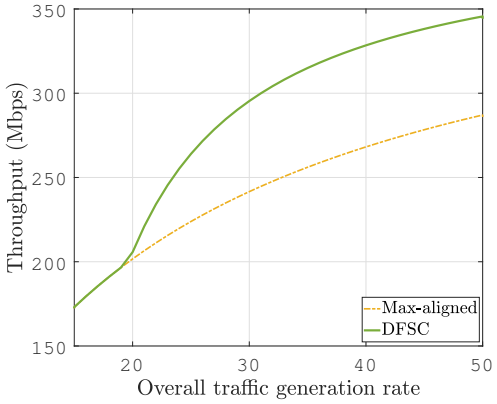


Fig. 4. Throughput when STAs have heterogeneous traffic generation rates in two-STA case

TABLE I
SIMULATION PARAMETERS (SINGLE BSS)

Parameter	Value
Channel bandwidth	160 MHz
Max number of RGs per STA	2
Number of RGs	4
Number of STAs	15
MCS (index)	[1 3 5]
Traffic generation rate	[10 40 70] Mbps
Max aggregation size	64 MPDUs
Slot length	9 us
SIFS	16 us
DIFS	34 us
Guard interval	400 ns (short)

according to Proposition 3. In the figure, N is the total number of user frames multiplexed in an MU transmission.

In the above design, service differentiation of different traffic classes (e.g. delay-sensitive traffic) is not considered. In order to serve different traffic classes better, a scheduler can give them different weights in calculating the fitness of a STA so that higher-priority class traffic is allocated more resources than others. In frame construction, a weight can also be given to a high-priority traffic class by allowing it to convey more data beyond the determined frame length if the class queue is not empty.

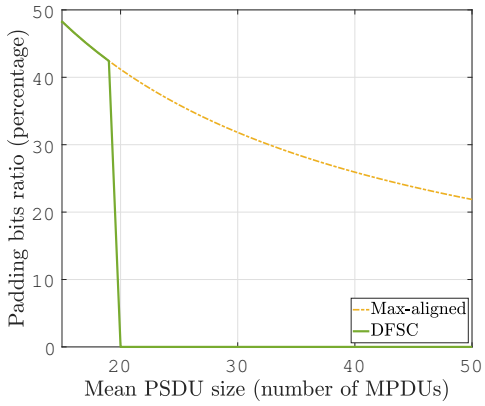


Fig. 5. Padding bits ratio of MU frames when STAs have heterogeneous traffic generation rates in two-STA case

VII. PERFORMANCE EVALUATION

In this section, we evaluate the effectiveness of DFSC in comparison with basic schemes for various network deployment scenarios, network parameters and scheduling algorithms. The Wi-Fi transmission considered in evaluation is based on IEEE 802.11ax [1].

We assume that link adaptation is performed perfectly. Therefore, a transmission failure results from collision only. We consider downlink traffic. The maximum number of MPDUs aggregated in A-MPDU is set to 64. The length of each MPDU is 1500 bytes. Each performance point is the average of ten simulation runs each of which is for one hundred seconds.[5]

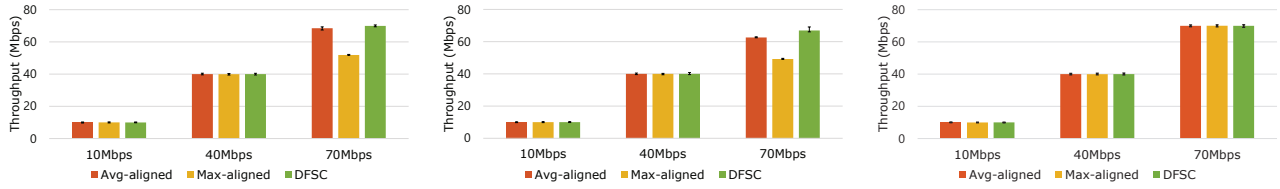
A. Two-Station Case

In order to get deeper understanding of the MU frame construction problem and the impact of padding bits based on numerical results, we first consider a simple case by varying the data sizes of two user frames and analyze the throughput performance. The transmission bit rate is fixed as 195 Mbps[6] for both STAs (thus denoted by r with no subscript).

First, we investigate how throughput is affected by various combinations of the data sizes of user frames and the results are shown in Fig. 3. In the figure, each curve is the total throughput for a varying data size of the second STA's user frame (l_2) with a fixed data size of the first STA's (l_1). We assume that l_2 is always equal to or larger than l_1; thus we have no sample points for $l_2 < l_1$ in the curves. From the figure, we see that for given l_1 the maximum throughput is achieved when l_2 is equal to l_1 and decreases as l_2 increases. This is because the increase of l_2 results in the increase of the gap between l_1 and l_2, thus requiring more padding bits for the first user frame. When l_1 is equal to rT_o, throughput is fixed to the transmission bit rate (195 Mbps) regardless of l_2, as pointed out in Section IV. The black solid line is the boundary of the achievable throughput region. If DFSC is applied, padding will be prevented by limiting l_2 such that

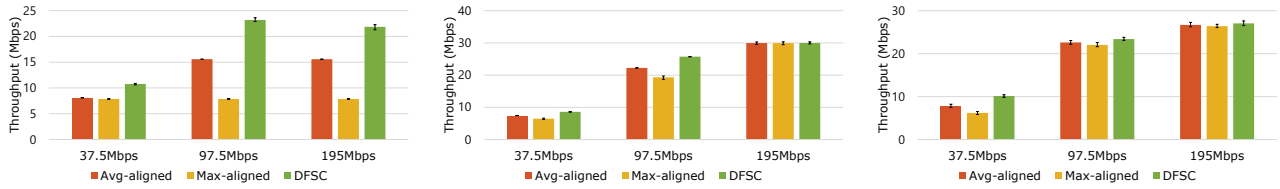
[5]We used a Matlab-based in-house simulator.

[6]It corresponds to MCS 8 with one 484-subcarrier RU and 1.6us guard interval in IEEE 802.11ax.



(a) RR (b) PF (c) QL

Fig. 6. Throughput comparison between different frame construction schemes for heterogeneous traffic generation rates and identical MCS rates among stations



(a) RR (b) PF (c) QL

Fig. 7. Throughput comparison between different frame construction schemes for identical traffic generation rates and heterogeneous MCS rates among stations

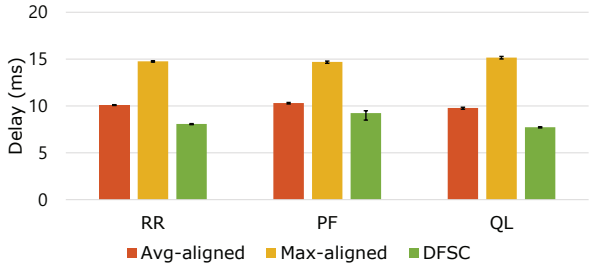


Fig. 8. Transmission delay for heterogeneous traffic generation rates and identical MCS rates among stations

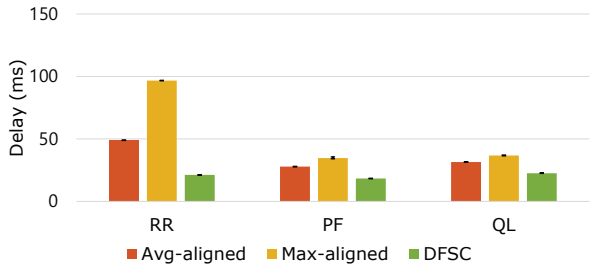


Fig. 9. Transmission delay for identical traffic generation rates and heterogeneous MCS rates among stations

user frames are constructed along the boundary line. Limiting l_2, however, will increase the amount of queued data for the second STA.

Next, we consider how the amount of generated traffic affects throughput and padding while two STAs have heterogeneous traffic demands. The results are shown in Fig. 4. A deterministic number of MPDUs are generated for each frame transmission as denoted by the *overall traffic generation rate* in the x axis of the figure; when it is 15, 1 MPDU is for STA 1 and the rest 14 MPDUs are for STA 2. The gap between the numbers of MPDUs for two STAs, i.e., 14, is fixed for all rates considered in the figure. If generated traffic is not fully included in the current transmission, the remaining amount is queued. As expected, DFSC outperforms Max-aligned due to the reduction of padding. However, DFSC behaves same as Max-aligned before the overall traffic generation rate gets 20 (3 for STA 1 and 17 for STA 2) which corresponds to $l_1 < rT_o$ ($\simeq$ 2.83 MPDUs) when all MPDUs for STA 1 are conveyed by the current frame; in this case, l_2 is not limited by DFSC (see Section IV). The reason behind this observation is is clearly seen in Fig. 5; the padding bits ratio of DFSC is same as Max-aligned in this range, thus implying that DFSC is not working (not limiting l_2) yet.

B. Single BSS Case

We now consider a BSS composed of an AP and 15 STAs connected to it. The channel bandwidth is 160 MHz. The total number of RGs is set to four, i.e., each RG corresponds to a RU occupying 40 MHz bandwidth in OFDMA of IEEE 802.11ax. We limit the maximum number of RGs allocated to a single STA by two. Three types of scheduling algorithms—round robin (RR), proportional fair (PF) and queue length-weighted (QL)—are considered. Packet arrivals follow the Poisson distribution. For a baseline, we consider the Max-aligned scheme. We consider two performance metrics:

- Throughput: It is defined as the amount of transmitted data in unit time. However, throughput is upper-limited by a traffic load and, if all STAs already achieve the throughput same as given traffic loads, we cannot see any difference between schemes. The following metric is needed in such a case.
- Delay: It is defined as the period from the time when a packet first arrives at a buffer to the time when it finally arrives at the destined node. It will show how fast queued

data is depleted and how efficient each transmission is. It is also highly related with the quality of service for delay-sensitive traffic and user experience. Although different schemes have the same throughput value, they may have different delay performance in unsaturated traffic conditions.

The results are shown as averaged values along with the minimum and maximum of per-user performance values for each averaged point as its confidence interval. The simulation parameters are summarized in Table I.

First, in order to consider heterogeneous traffic conditions of STAs, we divide them into three groups (low, medium and high traffic demands, respectively, each with five STAs) and set the traffic generation rates to 10 (low), 40 (medium) and 70 Mbps (high), respectively. The transmission bit rate of a STA when the STA is allocated a single RG, which we call modulation and coding scheme (MCS) rate in the paper, is set to 195 Mbps. The results are shown in Fig. 6. With RR (Fig. 6(a)) and PF (Fig. 6(b)), DFSC outperforms Max-aligned for the STAs with high traffic demands (70 Mbps). This is because only these STAs hunger for more resources in this scenario and DFSC has room to give them more transmission opportunities thanks to improved efficiency. Avg-aligned achieves higher efficiency than Max-aligned and thus reduces the performance gap with DFSC. With QL (Fig. 6(c)), no difference is observed between the schemes because the considered traffic demands are all handled well.

While the throughput gain of DFSC is marginal in the above scenario, the improvement of delay performance is significant as shown in Fig. 8. This is because throughput is upper-limited by a traffic generation rate and does not tell the difference between different schemes when generated traffic is handled all. The figure shows that DFSC reduces delay, especially with RR and QL scheduling by up to 50% since DFSC allows more transmission opportunities that become available from increased transmission efficiency and reduced airtime consumption.

Next, we consider heterogeneous MCS rates among STAs, i.e., STAs are in different channel conditions. For this, we divide the STAs into three groups (each with five STAs) and set the MCS rates to 32.5, 97.5 and 195 Mbps[7], respectively, for the groups. The traffic generation rate is now fixed for all STAs as 30 Mbps.

With RR (Fig. 7(a)), the Max-aligned scheme achieves the same throughput for all groups despite their heterogeneous MCS rates. Although a STA with a high MCS requires a short frame, other STAs with low MCSs make the resulting frame long with padding bits, thus reducing transmission efficiency. DFSC limits the frame length typically towards the STA with a good channel condition. As a result, the throughput of the STA using a higher MCS is increased while the throughput of the STA using the lower MCS is lowered. Avg-aligned performs better than Max-aligned, but not as good as DFSC because the length of a constructed frame is not optimized.

With PF (Fig. 7(b)), STAs with high MCS are given more frequent opportunities of transmission. Once a STA with low MCS is scheduled, it may have a large amount of queued data since it is typically not scheduled frequently. The Max-aligned scheme allows such a STA to transmit a large amount of data at a time and consume a long airtime in a transmission, thus resulting in a large amount of padding bits. DFSC restricts such a long airtime of a low-MCS STA and reduces padding bits, but at the expense of the reduced amount of transmission data at a time for low-MCS STAs. Thanks to the increased transmission efficiency and reduced overall airtime consumption, DFSC enables more transmission opportunities in the channel and increases the throughput of low-MCS STAs as well finally.

With QL (Fig. 7(c)), STAs with low MCS have increased throughput while those with high MCS have only a small increase of throughput over Avg and Max-aligned. This is because STAs with low MCS have longer queues than others and thus are prioritized by the scheduler.

The reduction of delay performance is more significant than the previous case as shown in Fig. 9. DFSC achieves 78% and 47% of reduction with RR and PF, respectively, compared to Max-aligned and 56% and 35% of reduction, respectively, compared to Avg-aligned. Even with QL, DFSC achieves 39% and 29% of reduction compared to Max and Avg-aligned schemes, respectively.

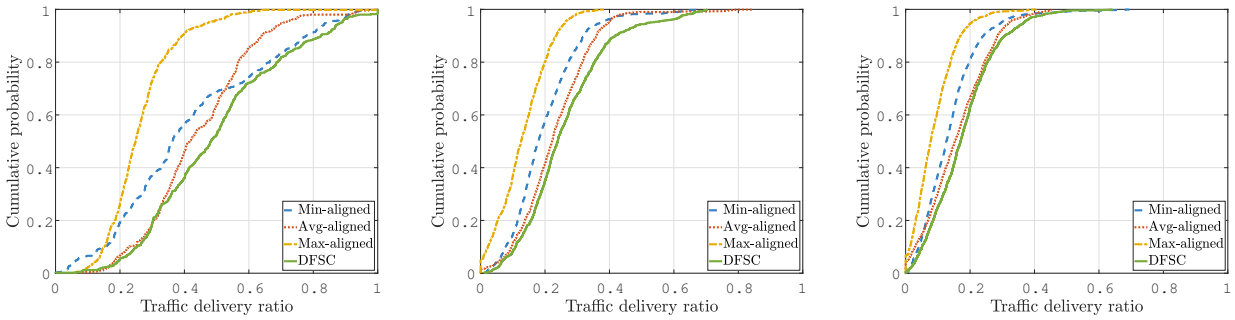
[7]These rates correspond to MCS indices 1, 4 and 8, respectively, with one 484-subcarrier RU and 1.6us guard interval in IEEE 802.11ax.

TABLE II
SIMULATION PARAMETERS (OVERLAPPING BSSS)

Parameter	Value
Channel bandwidth	80 MHz
Max number of RGs per STA	2
Number of RGs	4
Number of BSSs	[5 10 15]
Number of STAs per BSS	10
MCS (index)	[1 - 9]
Traffic generation rate	[0 - 25] Mbps
Max aggregation size	64 MPDUs
Slot length	9 us
SIFS	16 us
DIFS	34 us
Guard interval	400 ns (short)

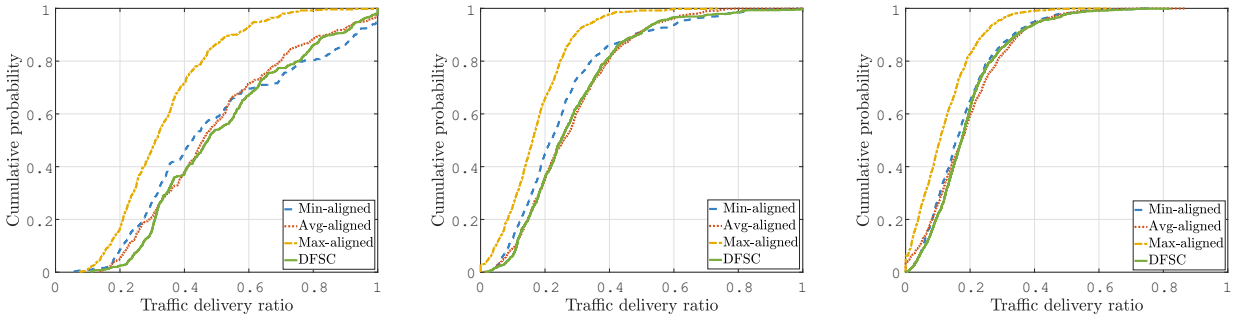
C. Overlapping BSSs Case

In this simulation scenario, we consider dense AP deployment cases where there exist multiple BSSs having overlapping coverage areas with each other. Each BSS is composed of an AP and 10 STAs connected to it. For each STA, a traffic generation rate is picked randomly in [0, 25]Mbps and MCS is selected in [1, 9] randomly with allocation of two RUs, thus ranging from 32.5 to 433.3 Mbps for each RU (with 484-subcarrier RU and 1.6us guard interval). Since STAs are given individual traffic generation rates and MCSs, we use *traffic delivery ratio* (TDR) to evaluate schemes, which is defined as the ratio of the delivered traffic amount to the generated amount; TDR of one means that all generated traffic is delivered to corresponding destinations in simulation time. The rest of the simulation parameters remain same as the single BSS case. The simulation parameters are summarized in Table II.



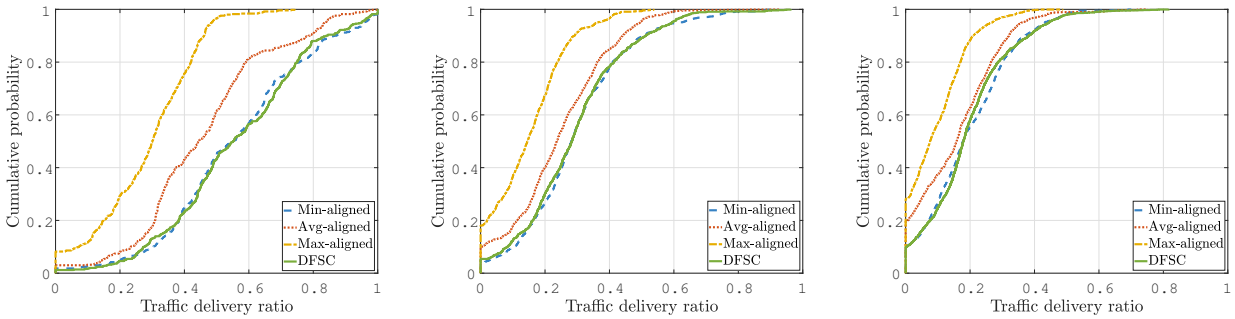
(a) 5 OBSSs (b) 10 OBSSs (c) 15 OBSSs

Fig. 10. CDFs of traffic delivery ratio with round robin scheduling



(a) 5 OBSSs (b) 10 OBSSs (c) 15 OBSSs

Fig. 11. CDFs of traffic delivery ratio with proportional fair scheduling

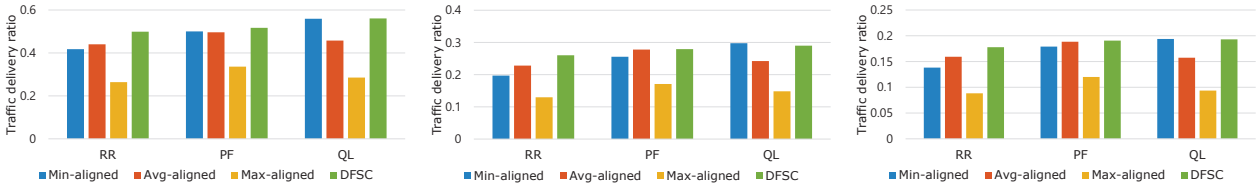


(a) 5 OBSSs (b) 10 OBSSs (c) 15 OBSSs

Fig. 12. CDFs of traffic delivery ratio with queue length-weighted scheduling

With RR (Fig. 10), both Min and Max-aligned show worse TDR performance than DFSC. Since RR gives equal opportunities to STAs even with small traffic, Min-aligned makes a frame too short while Max-aligned results in excessive padding bits. DFSC always achieves higher TDR than them. DFSC also outperforms Avg-aligned in all OBSSs cases. For 5 OBSSs, Avg-aligned achieves similar with DFSC for bottom 30% STAs, but DFSC achieves much higher TDR for the rest, thus leading to 14% improvement over Avg-aligned on average (20% and 91% improvement over Min and Max-aligned schemes, respectively). As the number of OBSSs increases, the basic three schemes achieve closer performance to each other's since STAs' queues pile up and all schemes tend to construct longer frames. The gain of DFSC over the best among the basic ones, i.e., Avg-aligned, is 15% for 10 OBSSs and 12% for 15 OBSSs.

Fig. 11 shows TDR performance with PF scheduling. As shown in the figure, the performance gap between DFSC and the other schemes is smaller than that with RR since PF scheduling itself has higher efficiency than RR. Another observation from the figure is the performance enhancement of the Min-aligned scheme; resource allocation to low-MCS STAs is less frequent with PF scheduling and thus too short frames are less made by Min-aligned. DFSC still shows the best TDR performance for most STAs in all OBSSs cases considered and the gain for average TDR over the basic ones



(a) 5 OBSSs (b) 10 OBSSs (c) 15 OBSSs

Fig. 13. Average traffic delivery ratio of various frame construction schemes combined with different scheduling algorithms in OBSSs environments

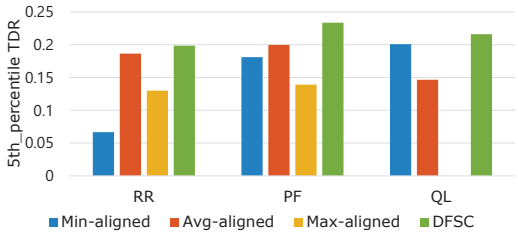


Fig. 14. 5th percentile of the traffic delivery ratio of various frame construction schemes combined with different scheduling algorithms in 5 OBSSs case

ranges from 1% (over Avg-aligned with 10 OBSSs) to 64% (over Max-aligned with 10 OBSSs).

The TDR results with QL scheduling are shown in Fig. 12. Since QL scheduling prioritizes STAs with more traffic in queue, it tends to produce long frames. For long frames, the overhead of padding bits dominates transmission efficiency and the protocol overhead has less impact. Thus Avg and Max-aligned schemes show lower performance than before and Min-aligned outperforms them due to minimization of padding bits. When the network is more crowded with 10 and 15 OBSSs, frames tend to have longer sizes due to increasing queues and Min-aligned even achieves almost equal TDR performance with DFSC.

Fig. 13 shows the summary of average TDR performance for all schemes under comparison. It is noted from the figure that the basic schemes achieve good performance in some cases, but not in all cases. For example, Avg-aligned performs similar with Min-aligned for 5 OBSSs, but gets better for 10 and 15 OBSSs; Min-aligned is good with QL scheduling, but gets worse than Avg-aligned with RR scheduling. However, DFSC achieves the best (or close to the best) performance among all schemes in all OBSSs cases with all scheduling algorithms, thus having wide adaptability to various network environments and scheduler combinations. In terms of 5th percentile TDR presented in Fig. 14, DFSC also outperforms all the other schemes with all scheduling algorithms.

VIII. Conclusion

We developed a data frame construction scheme for MU transmission in IEEE 802.11 WLANs, called DFSC. DFSC finds the optimal length of an MU frame to maximize its transmission efficiency, taking into account the buffer statuses and transmission bit rates of the stations which are allocated RG(s) in the frame. DFSC is combined with a general scheduling algorithm that selects a set of stations and allocates RG(s) to them for each frame transmission based on the algorithm-specific objective; the decision is then passed to DFSC and the final frame is constructed. Through comprehensive simulation, we demonstrated that the proposed scheme integrated with various scheduler algorithms outperforms basic schemes in a wide range of network environments.

References

[1] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications; Amendment 6: Enhancements for High Efficiency WLAN*, IEEE P802.11ax Draft 3.0, 2018.

[2] Y.-C. Hsu, K. C.-J. Lin, and W.-T. Chen, "Client-AP association for multiuser MIMO networks," in *Proc. IEEE ICC*, 2015, pp. 2154–2159.

[3] X. Xie and X. Zhang, "Scalable user selection for MU-MIMO networks," in *Proc. IEEE INFOCOM*, 2014, pp. 808–816.

[4] M. Esslaoui, F. Riera-Palou, and G. Femenias, "A fair MU-MIMO scheme for IEEE 802.11ac," in *Proc. IEEE ISWCS*, 2012, pp. 1049–1053.

[5] Z. Chen, X. Zhang, S. Wang, Y. Xu, J. Xiong, and X. Wang, "BUSH: Empowering large-scale MU-MIMO in WLANs with hybrid beamforming," in *Proc. IEEE INFOCOM*, 2017, pp. 1–9.

[6] D. Bankov, A. Didenko, E. Khorov, V. Loginov, and A. Lyakhov, "IEEE 802.11ax uplink scheduler to minimize delay: A classic problem with new constraints," in *Proc. IEEE PIMRC*, 2017, pp. 1–5.

[7] J. Choi, S. Choi, and K. B. Lee, "Sounding node set and sounding interval determination for IEEE 802.11ac MU-MIMO," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10 069–10 074, Mar. 2016.

[8] A. Zubow, "Downlink MIMO in IEEE 802.11ac-based infrastructure networks," in *Proc. IEEE GLOBECOM*, 2015, pp. 1–7.

[9] Y. Nomura, K. Mori, and H. Kobayashi, "Efficient frame aggregation with frame size adaptation for next generation MU-MIMO WLANs," in *Proc. IEEE NGMAST*, 2015, pp. 288–293.

[10] B. Bellalta, J. Barcelo, D. Staehle, A. Vinel, and M. Oliver, "On the performance of packet aggregation in IEEE 802.11ac MU-MIMO WLANs," *IEEE Commun. Lett.*, vol. 16, no. 10, pp. 1588–1591, Oct. 2012.

[11] A. U. Syed, M. Iftikhar, and L. Trajkovic, "Optimal PPDU duration algorithm for VHT MU-MIMO systems," in *Proc. IEEE ICCCN*, 2017, pp. 1–2.

[12] S. Sur, I. Pefkianakis, X. Zhang, and K.-H. Kim, "Practical MU-MIMO user selection on 802.11ac commodity networks," in *Proc. ACM MobiCom*, 2016, pp. 122–134.

[13] C.-H. Lin, Y.-T. Chen, K. C.-J. Lin, and W.-T. Chen, "acPad: Enhancing channel utilization for 802.11ac using packet padding," in *Proc. IEEE INFOCOM*, 2017, pp. 1–9.

[14] ——, "FDoF: Enhancing channel utilization for 802.11ac," *IEEE/ACM Trans. Netw.*, vol. 26, no. 1, pp. 465–477, Jan. 2018.

[15] N. Anand, J. Lee, S.-J. Lee, and E. W. Knightly, "Mode and user selection for multi-user MIMO WLANs without CSI," in *Proc. IEEE INFOCOM*, 2015, pp. 451–459.

[16] Y. Morino, T. Hiraguri, H. Yoshino, and K. Nishimori, "Proposal of overhead-less access control scheme for multi-beam massive MIMO transmission in WLAN systems," in *Proc. IEEE Med-Hoc-Net*, 2017, pp. 1–5.

[17] Z. Wu, X. Gao, and Y. Shi, "A novel MU-MIMO-OFDM scheme with the RBD precoding for the next generation WLAN," in *Proc. IEEE MILCOM*, 2015, pp. 565–569.

[18] M. Xie and T. M. Lok, "Access point selection and auction-based scheduling in uplink MU-MIMO WLANs," in *Proc. IEEE ICC*, 2016, pp. 1–6.
[19] Y. Zhou, A. Zhou, and M. Liu, "OUS: Optimal user selection in MU-MIMO WLAN," in *Proc. IEEE ICNC*, 2016, pp. 1–5.
[20] W. Lin, B. Li, M. Yang, Q. Qu, Z. Yan, X. Zuo, and B. Yang, "Integrated link-system level simulation platform for the next generation WLAN-IEEE 802.11ax," in *Proc. IEEE GLOBECOM*, 2016, pp. 1–7.
[21] H. Yang, D.-J. Deng, and K.-C. Chen, "Performance analysis of IEEE 802.11ax UL OFDMA-based random access mechanism," in *Proc. IEEE GLOBECOM*, 2017, pp. 1–6.
[22] L. Lanante, H. O. T. Uwai, Y. Nagao, M. Kurosaki, and C. Ghosh, "Performance analysis of the 802.11ax UL OFDMA random access protocol in dense networks," in *Proc. IEEE ICC*, 2017, pp. 1–6.
[23] R. P. F. Hoefel, "IEEE 802.11ax: Joint effects of power control and IQ imbalance mitigation schemes on the performance of OFDM uplink multi-user MIMO," in *Proc. IEEE VTC-Fall*, 2017, pp. 1–6.
[24] R. Liao, B. Bellalta, M. Oliver, and Z. Niu, "MU-MIMO MAC protocols for wireless local area networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 162–183, Dec. 2016.
[25] C. Zhu, C. Ngo, A. Bhatt, and Y. Kim, "Enhancing WLAN backoff procedures for downlink MU-MIMO support," in *Proc. IEEE WCNC*, 2013, pp. 368–373.
[26] A. Shahanaghi, A. Abbasfar, and K. M. Avval, "Fairness analysis of MAC protocols in MIMO networks using stochastic geometry," in *Proc. IEEE ANTS*, 2016, pp. 1–5.
[27] S. Shrestha, G. Fang, E. Dutkiewicz, and X. Huang, "Zeroforcing precoding based MAC design to address hidden terminals in MU-MIMO WLANs," in *Proc. IEEE ICT*, 2015, pp. 283–288.
[28] B.-S. Chen, K. C.-J. Lin, S.-L. Chiu, R. Lee, and H.-Y. Wei, "Multiplexing-diversity medium access for multi-user MIMO networks," *IEEE Trans. Mobile Comput.*, vol. 15, no. 5, pp. 1211–1223, Jun. 2016.
[29] M. Z. Ali, J. Misic, and V. B. Misic, "Access point controlled MAC (A-MAC) protocol for uplink multi-user transmission in IEEE 802.11ac," in *Proc. IEEE GLOBECOM*, 2017, pp. 1–7.
[30] M. Yazid, A. Ksentini, L. Bouallouche-Medjkoune, and D. Aïssani, "Enhancement of the TXOP sharing designed for DL-MU-MIMO IEEE 802.11ac WLANs," in *Proc. IEEE WCNC*, 2015, pp. 908–913.
[31] J. Cha, H. Jin, B. C. Jung, and D. K. Sung, "Performance comparison of downlink user multiplexing schemes in IEEE 802.11ac: Multi-user MIMO vs. frame aggregation," in *Proc. IEEE WCNC*, 2012, pp. 1514–1519.
[32] D. Nojima, L. J. Lanante, Y. Nagao, M. Kurosaki, and H. Ochi, "Performance evaluation for multi-user MIMO IEEE 802.11ac wireless LAN system," in *Proc. IEEE ICACT*, 2012, pp. 804–808.
[33] G. Redieteab, L. Cariou, P. Christin, and J.-F. Hlard, "SU/MU-MIMO in IEEE 802.11ac: PHY+MAC performance comparison for single antenna stations," in *Proc. IEEE WTS*, 2012, pp. 1–5.
[34] M. Abu-Tair and S. N. Bhatti, "IEEE 802.11ac MU-MIMO wireless LAN cells with legacy clients," in *Proc. IEEE AINA*, 2017, pp. 160–166.
[35] D. N. Thanh, G. T. Ha, and S. Shin, "Performance analysis of IEEE 802.11n CSMA/CA 4×4 multi-user MIMO in erroneous channel," in *Proc. IEEE ATC*, 2015, pp. 400–405.
[36] X. Ma, Q. Gao, J. Wang, V. Marojevic, and J. H. Reed, "Dynamic sounding for multi-user MIMO in wireless LANs," *IEEE Trans. Consum. Electron.*, vol. 63, no. 2, pp. 135–144, Aug. 2017.
[37] S. Shrestha, G. Fang, E. Dutkiewicz, and X. Huang, "Effect of CSI quantization on the average rate in MU-MIMO WLANs," in *Proc. IEEE CCNC*, 2016, pp. 824–828.
[38] M. O. Al-Kadri, A. Aijaz, and A. Nallanathan, "An energy-efficient full-duplex MAC protocol for distributed wireless networks," vol. 5, no. 1, pp. 44–47, Feb. 2016.
[39] A. Aijaz and P. Kulkarni, "Simultaneous transmit and receive operation in next generation IEEE 802.11 WLANs: A MAC protocol design approach," *IEEE Wireless Commun. Mag.*, vol. 24, no. 6, pp. 128–135, Dec. 2017.
[40] A. Tang and X. Wang, "A-duplex: Medium access control for efficient coexistence between full-duplex and half-duplex communications," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5871–5885, Oct. 2015.
[41] R. Liao, B. Bellalta, and M. Oliver, "Modelling and enhancing full-duplex MAC for single-hop 802.11 wireless networks," vol. 4, no. 4, pp. 349–352, Aug. 2015.
[42] S. Kim, M. S. Sim, C. B. Chae, and S. Choi, "Asymmetric simultaneous transmit and receive in WiFi networks," vol. 5, pp. 14 079–14 094, 2017.
[43] W. Choi, H. Lim, and A. Sabharwal, "Power-controlled medium access control protocol for full-duplex WiFi networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3601–3613, Jul. 2015.
[44] S. Y. Chen, T. F. Huang, K. C. J. Lin, Y. W. P. Hong, and A. Sabharwal, "Probabilistic medium access control for full-duplex networks with half-duplex clients," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2627–2640, Apr. 2017.
[45] J. Lee, H. Ahn, and C. Kim, "An OFDMA two-symbol coordination MAC protocol for full-duplex wireless networks," in *Proc. IEEE ICOIN*, 2017, pp. 344–348.
[46] G. Lee, H. Ahn, and C. Kim, "In-frame querying to utilize full duplex communication in IEEE 802.11ax," in *Proc. IEEE CSCN*, 2015, pp. 252–256.
[47] R. W. Freund and F. Jarre, "Solving the sum-of-ratios problem by an interior-point method," *Journal of Global Optimization*, vol. 19, pp. 83–102, 2001.

Sanghyun Kim is a Ph.D. student of the Department of Electrical and Information Engineering, Seoul National University of Science and Technology (SeoulTech), Seoul, Korea. His current research focuses on wireless networks in unlicensed spectrum.

Ji-Hoon Yun received the BS degree in electrical engineering from Seoul National University (SNU), Seoul, Korea, in 2000, and the MS and PhD degrees in electrical engineering and computer science from SNU, in 2002 and 2007, respectively.

He is currently an associate professor in the Department of Electrical and Information Engineering, Seoul National University of Science and Technology (SeoulTech), Seoul, Korea. Before joining SeoulTech in March 2012, he was in the Department of Computer Software Engineering, Kumoh National Institute of Technology, as an assistant professor. He was a postdoctoral researcher in the Real-Time Computing Laboratory, The University of Michigan, Ann Arbor, Michigan, in 2010 and a senior engineer in the Telecommunication Systems Division, Samsung Electronics, Suwon, Korea, from 2007 to 2009. His current research focuses on efficient computing of mobile devices and wireless networks.