

RESEARCH

Open Access

User scheduling for MU-MIMO transmission with active CSI feedback

Kyu-haeng Lee and Chong-kwon Kim*

Abstract

User scheduling boosts the multi-user multi-input multi-output (MU-MIMO) gain by selecting an optimal set of users to increase the 802.11 Wi-Fi system capacities. Many kinds of user scheduling algorithms, however, fail to fully realize the advantages of MU-MIMO due to considerable channel state information (CSI) overhead. In this paper, we propose a new MU-MIMO MAC protocol, called 802.11ac+, including a novel user scheduling algorithm. Unlike most proposals, where user scheduling is performed after an access point (AP) receives CSI from all users, 802.11ac+ determines the best user set during the CSI feedback phase. In particular, the AP broadcasts a channel hint about previously scheduled users, and the remaining users actively send CSI reports according to their effective channel gains (ECGs) calculated from the hint. Based on the proposed scheme, we develop two fair scheduling protocols, Round-Robin 802.11ac+ (RR-11ac+) and Proportional-Fair 802.11ac+ (PF-11ac+). Through trace-driven MATLAB simulations, we prove that the proposed schemes not only improve the throughput gain but also enhance the fairness among users.

1 Introduction

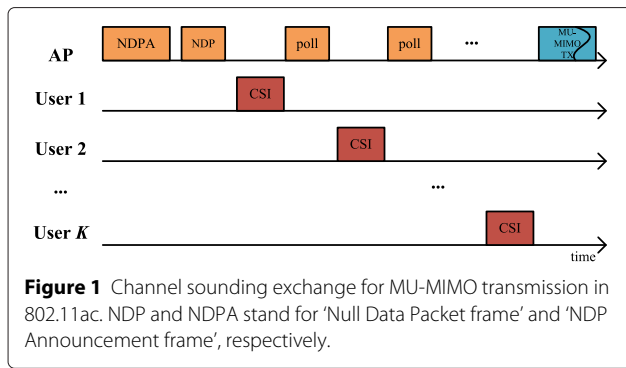
Most significant throughput gains of recent wireless communication systems come from multi-user multi-input multi-output (MU-MIMO) transmission using beamforming, which focuses energy toward a user so that the SNR and data rates increase. Wi-Fi standards as well as LTE-based cellular systems [1,2] have been adopting the MU-MIMO as a key technology for the next-generation wireless communication, and recently, several commercial products have been successfully deployed on the market.

The performance of MU-MIMO can be significantly improved when proper user scheduling is performed by multi-antenna nodes such as access points (APs) or base stations (BSs). Users with high received power or with an orthogonal channel can be chosen for increasing diversity gain or multiplexing gain, respectively. However, determining an optimal user set is nontrivial since it is generally integrated with finding optimal beamforming weights and powers. Even though we employ a simple beamforming strategy (e.g., zero-forcing beam forming (ZFBF)), a brute force search over all possible user sets is required and thus still creates a heavy burden. This motivates greedy

user selection algorithms aimed at providing sub-optimal performance with a practical low complexity [3-9].

To reap the benefit of the MU-MIMO user selection in the aforementioned algorithms, the *channel state information* (CSI) feedback overhead issue should be addressed. The downlink CSI of the candidate users must be efficiently fed back to the AP. In contrast to cellular systems [1] where separate control channels are used to report the CSI, current 802.11ac Wi-Fi systems use a series of poll-based CSI feedbacks for each user [2], as shown in Figure 1. Furthermore, the CSI feedback is transmitted at the low basic rate (e.g., 6.5 Mbps) and also grows as the number of transmitter/receiver antennas, quantization level, and subcarrier group size increases. The CSI feedback overhead can reach up to 25× compared to the data transmission time in the case of 160 MHz of bandwidth and 4×1 MIMO [10]. Such excessive overhead could easily overwhelm the multi-user diversity gain even under optimal user selection. Figure 2 shows the throughput gains of a user selection scheme over 802.11ac [2] as a function of the number of polls. As shown in the figure, throughput gain in the downlink decreases with the number of polls. In particular, when the AP accesses the CSI of all users, the loss increases to about 70%. Significant CSI overhead in the downlink also brings a slight throughput loss in the uplink, thereby degrading the whole system performance.

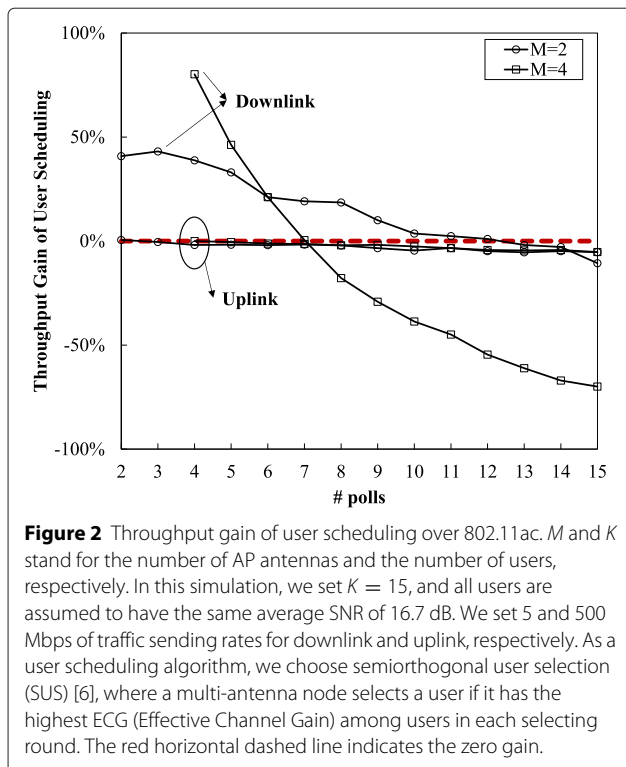
*Correspondence: ckim@snu.ac.kr
Department of CSE, Seoul National University, 151-744 Seoul, Korea



In this paper, we present an MU-MIMO MAC protocol for Wi-Fi systems, called 802.11ac+. In particular, 802.11ac+ provides a novel and practical user scheduling solution. Our idea is very intuitive: perform the user scheduling during the CSI feedback phase. To do this, the AP broadcasts channel information about previously scheduled users by appending it to a poll frame. We refer to this channel information as a *channel hint*. Then, users calculate their effective channel gains (ECGs) from the channel hint, and the user with the largest gain *actively* sends a CSI report back to the AP. Upon receiving the CSI report, the AP includes the user in the multi-user transmission schedule and repeats this process while its degrees of freedom (DoF) constraint is satisfied. As a result, the performance of 802.11ac+ is comparable to

other user scheduling heuristics, with the key difference that it requires a much smaller amount of CSI feedback.

The design of 802.11ac+ is challenging for the following reasons. First, the channel hint should be well designed to balance between overhead and efficiency. In order to reduce the overhead, the channel hint should be succinct, but at the same time, it should include all channel information of already scheduled users such that the remaining users can estimate their ECGs. To deal with such a trade-off, we propose an efficient channel hint broadcasting mechanism. More specifically, an AP extracts the effective channel vector of the last scheduled user from the received CSI reports and uses it as a channel hint. The second challenging issue comes from the fact that users hardly know how much better their ECGs are in a fully distributed manner, which leads us to adopt a contention for users to get a feedback opportunity. Due to the nature of contention, there may be a frame collision if more than one user sends a CSI report at the same time. To address the collision problem, we use a delayed feedback approach, where users delay their CSI feedback according to their ECGs: the bigger the ECG a user has, the faster the CSI report will be. To improve the performance, we focus on maximizing the probability of feedback success. Additionally, we extend our user selection algorithm to well-known fair scheduling protocols, Round-Robin (RR) and Proportional-Fair (PF), to achieve fairness among users.



To evaluate the performance of our approach, we implement the 802.11ac+ and its fair scheduling protocols on a MATLAB simulator. The extensive trace-driven simulation results show that 802.11ac+ obtains much higher throughput gain than 802.11ac and a MAC protocol employing a well-known user scheduling algorithm. Also, we demonstrate that two fair scheduling protocols of 802.11ac+ give a much better throughput fairness than 802.11ac, especially when users experience different channel qualities.

We summarize our main contributions as follows. First, we propose the design of an MU-MIMO MAC protocol, called 802.11ac+, which accomplishes user scheduling gain with a far smaller amount of CSI feedback. Second, in order to realize the protocol, we provide a channel hint broadcasting mechanism and an active CSI feedback scheme. Third, we build two fair scheduling protocols based on the user scheduling of 802.11ac+ to achieve fairness among users. Lastly, the performance of 802.11ac+ has been evaluated through trace-driven simulations in MATLAB.

The remainder of the paper is organized as follows. Section 2 provides the related work, and we next describe the 802.11ac+ mechanism in greater detail in Section 3. In Section 4, we discuss the fair scheduling protocols of

802.11ac+, and Section 5 shows the performance evaluation. We finally conclude our paper in Section 6.

2 Related work

We survey the research results on the user scheduling schemes and channel feedback overhead reduction techniques related to 802.11ac+.

2.1 User scheduling schemes

User selection has been highlighted since it can improve the MU-MIMO performance significantly. In user scheduling, it is challenging and often impractical to determine the optimal user set due to the large search space. Therefore, some protocols based on heuristics have been proposed. Many publications study the problem of maximizing sum capacity [3,7-9]. SUS [6] selects one user in each round by exploiting channel orthogonality between users. However, the selected users cannot guarantee that they always increase the sum rate. GUSS [7] considers ‘delete’ and ‘swap’ operations to guarantee a positive increment of channel capacity in each selecting round. Jin et al. propose a volume metric as the product of diagonal elements of an upper-triangular matrix by performing QR factorization to the selected user channels [8]. Shen et al. [9] propose and compare two algorithms for both approaches with block diagonalization [11] which is a generalized concept of channel inversion.

In addition to the sum rate maximization, the rate balancing problem aimed at maximizing the throughput subject to the constraint that the rates of the different users need to have certain fixed ratios has also been considered in many literatures [12,13]. Hellings et al. propose a gradient projection-based solution [12], and Guthy et al. propose a user classification scheme based on perturbation analysis [13], in order to handle the rate balancing problem. Along this line, Lima et al. consider a frequency allocation problem for MIMO-OFDMA networks [14]. Unfortunately, these schemes come with a nonnegligible performance loss due to the excessive feedback overhead.

2.2 Channel feedback overhead reduction techniques

One way of overcoming the CSI overhead is to use the compression techniques for reducing CSI feedback bits. Codebook and quantization are already adopted in LTE and MIMO-based Wi-Fi systems [1,2,15-17]; however, selecting the optimal quantization level is still an open problem. In Wi-Fi systems, it is proposed to compress the CSI report along three dimensions: time, frequency, and quantization level. Even though compression is used, it is reported that the feedback can take about 25 times longer than the data transmission time when sending a small packet with a high data rate [10]. A vast literature of work has aimed to design an efficient CSI quantization mechanism [18-22]. Recently, compressive sensing has also been

used for feedback reduction in MIMO communications [23,24]. The downside of compressions is, as expected, a throughput loss of MU-MIMO transmission: fewer bits may offer diminishing returns.

In addition to the compression, the reduction of feedback loads by adaptively sending feedback can be used. One possible solution is to allow only users whose signal quality (e.g., SNR, carrier-to-noise ratio (CNR)) is higher than a pre-defined threshold to report their CSI [10,25-28]. However, these approaches may fail to obtain higher effective channel gain because their CSI reports are sent without considering the relationship between user channels.

3 802.11ac+

3.1 System model

In this paper, we consider a single basic service set (BSS) Wi-Fi network, where an M -antenna AP and K single-antenna user stations communicate with each other, as shown in Figure 3. We assume that MIMO channels satisfy the i.i.d. Rayleigh fading condition: the real and imaginary components of \mathbf{h}_{km} are i.i.d. Gaussian with unitary variance and zero mean, where \mathbf{h}_{km} is a channel gain from the m th antenna of the AP to the k th user. Then, we can characterize the channel of user k from the AP, i.e., \mathbf{h}_k , as a zero mean complex Gaussian channel vector.

In MU-MIMO, the multi-antenna transmitter, e.g., AP, uses precoding to send multiple data streams among several users at a given instant. Similar to most recent schemes, we incorporate zero-forcing beamforming (ZFBF) as the precoding strategy, since it effectively removes the mutual interference among concurrent transmissions by using a low-complexity precoding matrix computation. In ZFBF, the precoding vector of one user is selected to be orthogonal to the channel vector of the other users. Let S be the group of selected users. Then, the precoding matrix for S , denoted by $W(S)$, is obtained as:

$$W(S) = H(S)^\dagger = H(S)^*(H(S)H(S)^*)^{-1} \quad (1)$$

where $(\cdot)^\dagger$, $H(S)$, and H^* stand for a pseudo-inverse, the channel matrix of S , and the conjugate transpose of H , respectively.

3.2 Overview

As explained earlier, we do not separate CSI feedback from the user scheduling procedure. This is the key difference between the user scheduling in 802.11ac+ and other approaches. Figure 4 illustrates the main idea. At the beginning, an AP polls the first user (i.e., user 1 in this example) and it replies with its CSI. Then, the AP immediately joins user 1 to the scheduled user set S . We note that the first user is selected by the AP’s queuing policy. By using the first CSI, the AP makes a channel hint, and it broadcasts a polling frame with the channel hint to all

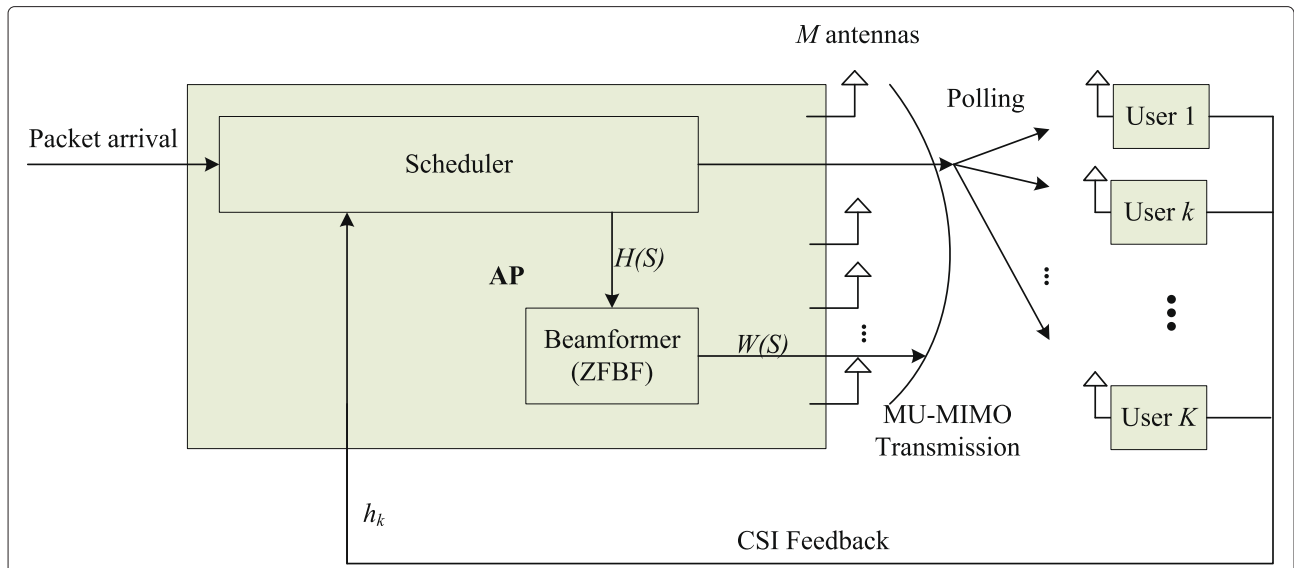


Figure 3 MU-MIMO downlink system with the M -antenna AP and K single-antenna user stations.

users. Upon receiving the poll, users compute their ECGs from the channel hint. Then, the user with the highest gain sends its CSI report to the AP through a contention. In this example, user 2 has the largest channel gain so it can feed back CSI for the second time. This step repeats until the AP successfully receives M CSI reports or the feedback timeout is triggered.

The intuition behind our user scheduling method is that users actively participate in scheduling decisions, unlike in most user scheduling algorithms, which are executed at only the multi-antenna transmitters. As a result, user scheduling in 802.11ac+ can limit the number of sounding exchanges to the number of AP antennas, while still keeping the scheduling gain. In the next section, we introduce two main mechanisms to realize 802.11ac+: channel hint broadcasting and active CSI feedback.

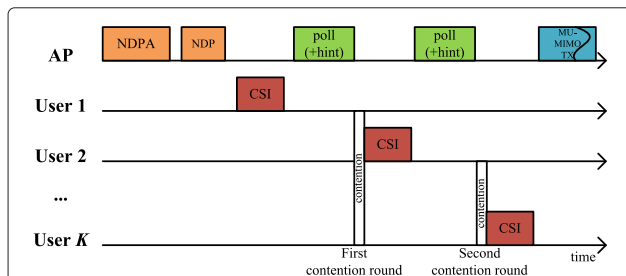


Figure 4 Operation example of 802.11ac+. Both AP and users participate in scheduling decisions. Except for the first scheduled user, all users should send a feedback report through a contention. If the AP fails in reception of any CSI reports, then it does not go to the next contention round and finishes the feedback procedure. Therefore, there could be maximum $(M - 1)$ contention rounds in the case of M AP antennas.

3.3 Channel hint broadcasting

3.3.1 Effective channel gain

Since the user scheduling in 802.11ac+ is based on harnessing the ECG, we first look into the concept of it. For a user channel \mathbf{h}_k , its effective channel vector^a is defined as \mathbf{e}_k and can be calculated by projecting \mathbf{h}_k onto the orthogonal complement of the subspace spanned by set $\{\mathbf{e}_{(1)}, \dots, \mathbf{e}_{(|S|)}\}$, being interpreted as the effective channel vectors of previously selected users [6,7]:

$$\mathbf{e}_k = \mathbf{h}_k - \sum_{j=1}^{|S|} \frac{\mathbf{h}_k \mathbf{e}_{(j)}^*}{\|\mathbf{e}_{(j)}\|^2} \mathbf{e}_{(j)} \quad (2)$$

$$= \mathbf{h}_k \left(I - \sum_{j=1}^{|S|} \frac{\mathbf{e}_{(j)}^* \mathbf{e}_{(j)}}{\|\mathbf{e}_{(j)}\|^2} \right) \quad (3)$$

$$= \mathbf{h}_k Q(S) \quad (4)$$

where I is the identity matrix and $Q(S)$ is a projection matrix. Then, the ECG of user k is denoted as $\|\mathbf{e}_k\|^2$.

Specifically, for OFDM systems, which divide the bandwidth into orthogonal subcarriers and treat each of the subcarriers as an independent narrowband channel, the ECG should be averaged over all subcarriers^b. Let $\mathbf{e}_k[c]$ be the ECG on subcarrier c , then we have:

$$\|\mathbf{e}_k\|^2 = \frac{1}{N_c} \sum_{c=1}^{N_c} \|\mathbf{e}_k[c]\|^2 \quad (5)$$

where N_c is the total number of subcarriers.

3.3.2 Channel hint design

There are many ways to implement a channel hint. In a simple way, the AP can use a projection matrix as a

channel hint. Although this simplifies the computation burden at the receiver, it always consumes significant bits for representing an $M \times M$ matrix per each subcarrier (the maximum required number of appended channel vectors is $(M - 1)MN_c$). By contrast, if we use the received CSI as a channel hint, then this will increase the receiver side computing complexity.

Our proposed design exploits a compromise between the above two schemes. Every time the AP receives a CSI report, it computes the effective channel vector from that CSI and uses it as a channel hint. Hence, in every poll frame, the effective channel vector of the last scheduled user will be included. On the receiver side, users should maintain the value of summation in Equation 3 during the feedback phase so that they can update a projection matrix and obtain their ECGs much faster.

We illustrate the poll frame structure used in 802.11ac+ in Figure 5. As shown in the figure, the poll frame requires only one additional channel vector, and thus, the maximum number of appended vectors is $(M - 1)N_c$. Since in each selection round, each user and the AP require one $(1 \times M) \times (M \times M)$ vector-matrix multiplication per subcarrier, we conclude that the computational complexity is $C(M - 1)N_c$, where C is a computational cost corresponding to one vector-matrix multiplication.

One may claim that polling additional $(M - 1)$ users and taking CSI from them for user selection would be more effective than using $(M - 1)$ channel hints. However, this may have a large feedback overhead, which brings MU-MIMO performance degradation. To clarify this, we compare the overhead of two schemes. Let τ_{ECV} , τ_{POLL} , and τ_{CSI} be the transmission time of sending one effective channel vector in channel hint, one polling frame, and one CSI feedback, respectively. Note that $\tau_{ECV} + \tau_{POLL}$ is the transmission time for one channel hint, and $\tau_{ECV} < \tau_{CSI}$. First, in 802.11ac+, the AP will send maximum $(M - 1)$ channel hints, and maximum M users will report their CSI reports. Then, we have the following:

$$\begin{aligned} \text{overhead}_{802.11ac+} &= \tau_{POLL} + (M - 1)(\tau_{ECV} + \tau_{POLL}) \\ &\quad + M\tau_{CSI} + 2M\tau_{SIFS} + (M - 1)\tau_{contention} \end{aligned} \tag{6}$$

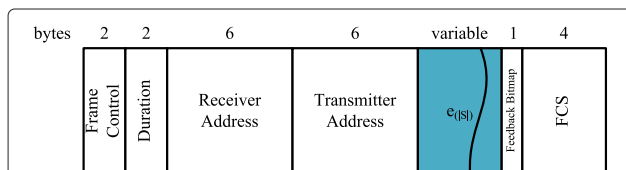


Figure 5 The change in use of the polling frame. When using a channel hint, polling is no longer destined for only one target user, so the receiver address should be the broadcast address. The length of the channel hint varies depending on the quantization level.

$$\begin{aligned} &= M\tau_{POLL} + (M - 1)\tau_{ECV} + M\tau_{CSI} + 2M\tau_{SIFS} \\ &\quad + (M - 1)\tau_{contention} \end{aligned} \tag{7}$$

where τ_{SIFS} stands for 802.11 Short InterFrame Space, which is $16 \mu s$ in 802.11ac [2]. Here, $\tau_{contention}$ is an overhead from the feedback contention of 802.11ac+, which will be discussed in the next section. As will be discussed later, the maximum value of a $\tau_{contention}$ is about $2\tau_{SIFS}$ (four slots per each round).

Now, assume that the AP polls $(M - 1)$ users after taking CSI from M users. In other words, the AP takes CSI from total $(2M - 1)$ users for user selection. Then, we have the following:

$$\begin{aligned} \text{overhead}_{add} &= (2M - 1)(\tau_{POLL} + \tau_{CSI}) + 2(2M - 1)\tau_{SIFS} \tag{8} \\ &= (2M - 1)\tau_{POLL} + (2M - 1)\tau_{CSI} \\ &\quad + 2(2M - 1)\tau_{SIFS}. \end{aligned} \tag{9}$$

Even if taking $\tau_{ECV} \approx \tau_{CSI}$, $\text{overhead}_{802.11ac+} < \text{overhead}_{add}$. Also, since the AP has no choice but to poll users randomly in this scheme, the user selection gain could be small, as shown in the case of $M = 4$ in Figure 2.

When sending a channel hint, since the AP may not intend all users to participate in the feedback procedure, it requires a method to notify only some dedicated users. For example, an AP should prevent users from sending CSI reports when it has no frames for them. For target user notification, we utilize the legacy Null Data Packet Announcement (NDPA) frame, by using fewer bits for the Association ID in the STA Info field and mapping bit positions to each user ID. This method maintains the same frame format and thus is efficient. A detailed description of the notification is out of the scope of this paper.

3.4 Active CSI feedback

3.4.1 Feedback contention

In 802.11ac+, only the user with the highest ECG should respond to the poll. However, this is a challenging task since users cannot know the ECG of the others in a distributed manner. If two users, who might think their gains are the best, feed back their CSI reports simultaneously, a frame collision will occur. Without any coordination between users, they should get opportunities to send feedback through a contention like the 802.11 distributed coordinate function (DCF). We call this a *feedback contention* in this paper.

In order to resolve the feedback contention, we adopt a delayed transmission approach. In this approach, users delay their feedback transmission regarding their ECGs, such that a user with higher value can access the feedback opportunity faster. To realize this, we first apply slotted time in our scheme, where the slot duration is fixed and each slot has a pre-defined threshold. We refer to this threshold as a *slot threshold*. A user is assigned to the specific slot according to its ECG and is allowed to transmit

feedback only at the slot. If a user senses a feedback transmission earlier than its slot, it gives up and waits until the next poll.

Since users should listen to the medium until they obtain a transmission opportunity, the minimum duration of the slot requires at least one slot time of 802.11 (e.g., 4 μ s for CCA + 5 μ s for RX/TX turnaround). To limit the expense of delayed feedback, an AP uses the timeout threshold (γ). If the AP does not receive any feedback reports before γ , then, it finishes the scheduling. When feedback collision or timeout occurs in the first contention round, the scheduling may be stopped with only one user (i.e., the first user) remaining in S . In this case, the AP just sends a frame to that user in diversity mode, which guarantees the minimal level of system performance.

Figure 6 shows an example of an active feedback scheme with three users and five feedback slots for the first contention round. The associated slot thresholds are given in Table 1, which will be explained in the following section. Assume that user 2 and user 1 can transmit CSI feedback in the second and fourth slots, respectively, while user 3 cannot get an opportunity due to the low gain. User 2 sends the CSI feedback in the second slot. The others listen to the feedback transmission and wait for the next contention round.

3.4.2 Slot threshold optimization

The performance of the active feedback mainly relies on how to select the thresholds: γ and slot thresholds. As described before, active feedback fails when a collision or timeout occurs; however, their impacts are different from each other. The loss from collision is much bigger than that from timeout, due to the relatively long time wasted in the collision. Thus, we first fix γ and find optimal values of the slot thresholds for each contention round. Now, we describe how to determine the thresholds as below.

Provided that a random channel vector on subcarrier c , $\mathbf{h}_k[c]$, is a zero-mean circularly symmetric complex Gaussian random variable, from Equation 5, $\|\mathbf{e}_k\|^2$ is a Gamma-distributed random variable with $N_c L$ (shape

Table 1 An example of the thresholds

CR	(w_s, w_c, w_t)	α_1	α_2	α_3	α_4	α_5
1	(1,1,1)	3.667	3.543	3.453	3.371	3.280
	(0.4,0.4,0.2)	3.680	3.558	3.472	3.396	3.318
2	(1,1,1)	2.541	2.436	2.360	2.291	2.215
	(0.4,0.4,0.2)	2.552	2.449	2.376	2.312	2.246
3	(1,1,1)	1.384	1.304	1.247	1.196	1.140
	(0.4,0.4,0.2)	1.392	1.314	1.259	1.212	1.163

parameter) and $1/N_c$ (scale parameter), where $L = \text{rank}(Q(S))$. Formally, for a constant α , $\|\mathbf{e}_k\|^2$ satisfies:

$$P\{\|\mathbf{e}_k\|^2 \leq \alpha\} = F(\alpha; N_c L, 1/N_c) \tag{10}$$

where $F(\alpha; N_c L, 1/N_c)$ is the cumulative distribution function (CDF) of the Gamma distribution for α , N_c and L . For fixed N_c and L , we use $F(\alpha)$ and $F(\alpha; N_c L, 1/N_c)$ in the same sense. Note that L varies every contention round.

For a timeout γ ($= G$ slots) in a certain contention round, let $\alpha_1, \alpha_2, \dots, \alpha_G$ be the threshold corresponding to each slot, where $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_G \geq 0$. We assume that γ is fixed over all contention rounds. We define the probability of successful feedback as the probability that only one user will send a CSI report in γ . Then, it can be calculated by summing up all success probabilities over all time slots:

$$\begin{aligned}
 p^{\text{suc}} &= \sum_{g=1}^G p_g^{\text{suc}} \tag{11} \\
 &= K' \{1 - F(\alpha_1)\} F(\alpha_1)^{(K'-1)} \\
 &\quad + K' \{F(\alpha_1) - F(\alpha_2)\} F(\alpha_2)^{(K'-1)} \\
 &\quad + \dots \\
 &\quad + K' \{F(\alpha_{G-1}) - F(\alpha_G)\} F(\alpha_G)^{(K'-1)} \tag{12}
 \end{aligned}$$

where K' is the number of contending users in the contention round. We note that $K' \leq K - 1$ because the first user is already selected by AP, and K' can be estimated by using the target user notification.

The timeout and collision probabilities are defined as follows:

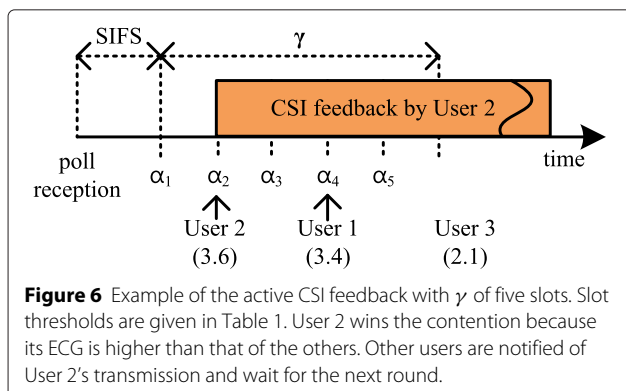
$$p^{\text{tout}} = F(\alpha_G)^{K'} \tag{13}$$

$$p^{\text{col}} = 1 - p^{\text{suc}} - p^{\text{tout}}. \tag{14}$$

Then, our objective is to find a set of thresholds as follows:

$$\text{argmax}_{\alpha_1, \dots, \alpha_G} w_s p^{\text{suc}} - w_c p^{\text{col}} - w_t p^{\text{tot}} \tag{15}$$

where w_s, w_c , and w_t are weights for the probabilities of success, collision, and timeout, respectively.



The optimization problem described in Equation 15 is nonconvex and is solved using nonlinear optimization technique. Note that it only needs to be solved whenever the network configuration (N_c, M, K) is changed or we can construct a set of solutions offline. Once a set of thresholds has been constructed, users can refer to them at any time.

Table 1 shows an example of the thresholds in the case of $\gamma = 5, M = 4, K' = 14$, and $N_c = 30$, where CR stands for contention round. Recall that there are maximum $(M - 1)$ contention rounds, and each round consists of γ time slots. From the table, we can see that the thresholds decrease during the course of a contention round because successful feedback reduces L one by one and hence decreases the DoF of the chi-square distribution. In addition, the reduced number of contending users also results in lower thresholds in further contention rounds. The thresholds of nonequal weights are slightly bigger than those of equal weights. In this case, the probability of success as well as the probability of collision decreases.

We compare the downlink throughput and success feedback probability of 802.11ac+ as a function of γ in Figure 7. In particular, to observe the impact of the threshold optimization according to channel distributions, we use two different channel distributions. Figure 7a shows the ECG distributions of two traces used in this simulation: ‘real trace’ and ‘synthetic trace (Rayleigh)’. Note that the slot thresholds are computed from the real trace. As shown in Figure 7b, a throughput loss (maximum 4 Mbps when $\gamma = 4$) happens in the synthetic trace scenario. This result is predictable: the thresholds are optimized only for the real trace. Even though the performance loss might seem small in this result, using a more exact channel model will improve the robustness of the proposed scheme. Note that in this section, we provide an analysis of the Rayleigh fading case only, but we can apply the same technique for different distributions, e.g., Rician, by changing Equation 10 in the optimization.

The graph also demonstrates that nonequal weights reduce both the feedback success probability and the collision probability (omitted here), as expected, thereby decreasing the scheduling gain. As a result, the throughput of nonequal weights is lower than that of the equal weights over all cases. Additionally, we can see that too large a γ decreases the effective channel gain while too small a γ decreases the user diversity gain. For the real trace scenario, both downlink throughputs increase as γ reaches four; after that, they decrease (for the synthetic trace, maximum downlink throughput is obtained when γ is three).

4 Fair scheduling

In this section, we develop two fair scheduling approaches of the 802.11ac+ based on the Round-Robin and

Proportional-Fair algorithms. Two main changes are adopted in the fair scheduling protocols to maintain the property of the 802.11ac+ user scheduling. The first change is to make use of a simple fair queue in the AP of both RR-11ac+ and PF-11ac+. In particular, the AP selects the first user by using the following:

$$\operatorname{argmax}_{k \in A} \frac{1}{\mathcal{R}_k} \quad (16)$$

where \mathcal{R}_k is the average data rate of user k and A is a target user set. For the remaining users, each fair scheduling protocol uses a different scheduling strategy, but the main idea is the same as that of 802.11ac+. We describe the details of RR-11ac+ and PF-11ac+ in the next subsections.

4.1 RR-11ac+

This is the simplest way to obtain the fairness that gives users equal transmission opportunities. Most Round-Robin schedulers use the average data rates of all users, and they can be monitored at the scheduler, which is very easy to implement.

In RR-11ac+, we use the same metric to schedule users, but a simple change is made for multi-user transmission: before the active CSI feedback, an AP limits the contending users according to their average data rates. Specifically, among remaining users in A , the AP selects top $(M - 1 + o)$ users in ascending order of average data rates such that users who have been served less have more feedback opportunities. If the nonnegative parameter o is set to 0, then RR-11ac+ is reduced to the 802.11ac with the legacy Round-Robin scheduling; however, it still considers the ECG of users. For example, if one’s ECG is too low so joining it into S is no longer helpful for multi-user transmission, then it will be discarded.

4.2 PF-11ac+

The proportional fairness maintains a balance between maximizing the sum rate and allowing all users at least a minimal level of service. Basically, both the original Proportional-Fair algorithms and PF-11ac+ use the following criterion to select users:

$$\frac{\mathcal{T}_k}{\mathcal{R}_k} \quad (17)$$

where \mathcal{T}_k is the current available data rate of user k .

The difficulty of legacy Proportional-Fair scheduling is that the scheduler cannot know the exact data rates of users (\mathcal{T}_k) when the scheduling decision has to be made. However, the active feedback of 802.11ac+ provides a simple implementation of the Proportional-Fair scheduling since users can actively join the scheduling decision. In particular, they can estimate \mathcal{T}_k by using their ECGs (i.e., $\mathcal{T}_k \sim \log(\|\mathbf{e}_k\|^2)$). Note that in PF-11ac+, users should maintain their average data rates, and the slot thresholds

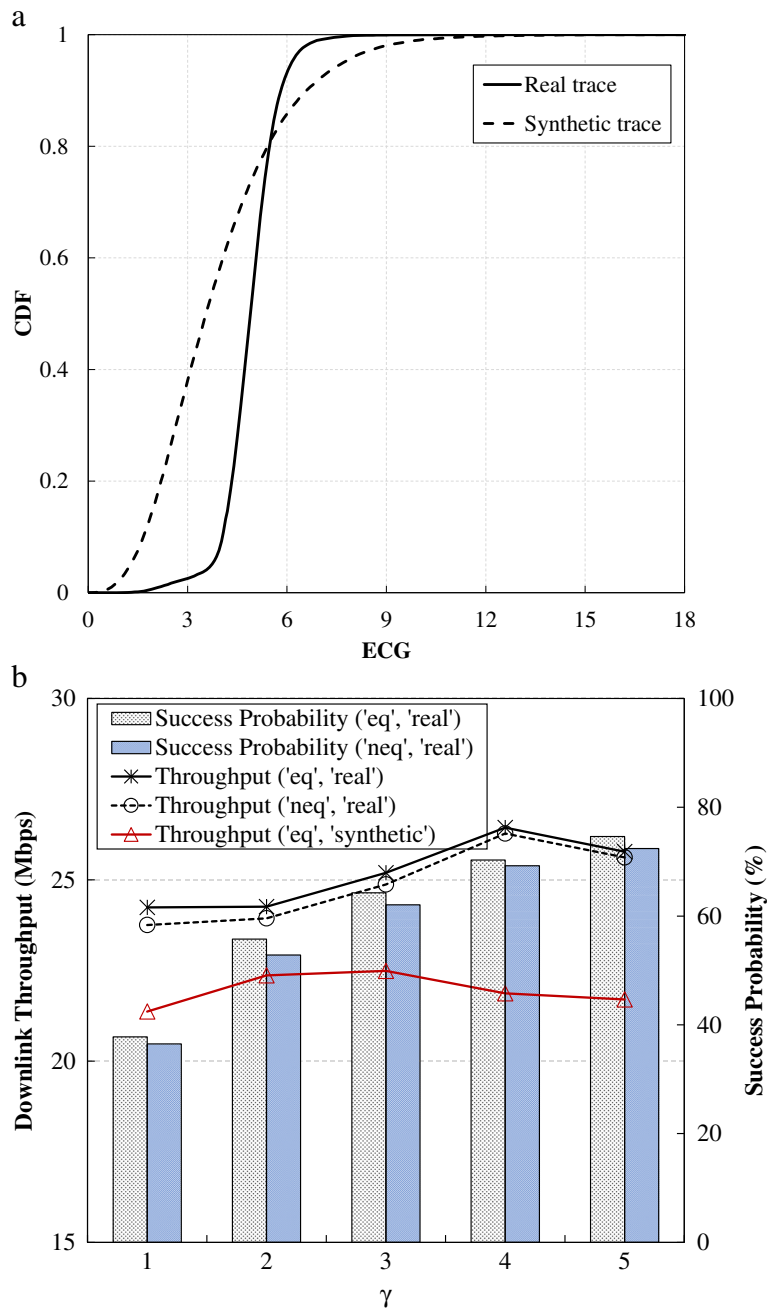


Figure 7 Performance of the active CSI feedback. **(a)** We use two different CSI traces (real and synthetic traces). The channel coefficients of the synthetic trace follow the Rayleigh distribution explained in Section 3.1. **(b)** We use weights of (0.4,0.4,0.2) as the nonequal weights cases. The results for ('neq', 'synthetic') are omitted because they show a similar pattern as those for 'real'. **(a)** ECG distributions. **(b)** Performance comparison with respect to γ .

should be redefined according to Equation 17 by adopting the technique in Section 3.

4.3 Summary

Table 2 shows the comparison between 802.11ac+, RR-11ac+, and PF-11ac+. As shown in the table, there are notable differences in three aspects: target user set, first

user selection, and n th user selection. First, in 802.11ac+ and PF-11ac+, all users are considered for scheduling, while in RR-11ac+, the number of target users is bounded to $M - 1 + o$, as explained earlier. Second, in 802.11ac+, basically, the first user is selected randomly according to AP's queuing policy, while in RR-11ac+ and PF-11ac+, the user with maximum utility Equation 16 will be chosen.

Table 2 Comparison between three schemes

	Size of target user set	First user selection	Metric for $n(> 1)$ th user selection
802.11ac+	K	random or $\operatorname{argmax}_k \ \mathbf{e}_k\ ^2$	$\ \mathbf{e}_k\ ^2$
RR-11ac+	$M - 1 + o$	$\operatorname{argmax}_k \frac{1}{\mathcal{R}_k}$	$\frac{\mathcal{T}_k}{\mathcal{R}_k}$
PF-11ac+	K		

To maximize the sum capacity, the user with the largest ECG will be selected as a first user, similar to SUS [6]. In using this method, one possible concern is that optimally selecting the first user is difficult since no CSI is available at the time of user selection. To handle this issue, we propose referring to the statistics of previous channel status of the users, for the first user selection. This method should work well when transmissions of the AP happen in channel coherent time, which typically ranges from 15 to 100 ms [10,29]. Lastly, for n th user selection, 802.11ac+ uses ECG as a selection metric, while the fair scheduling protocols use Equation 17 as a selection metric.

5 Performance evaluation

In this section, simulation results for the throughput and fairness performance of the proposed schemes are presented.

5.1 Setting

We implement 802.11ac, an SUS [6] based MAC protocol, 802.11ac+, RR-11ac+, and PF-11ac+ on the MATLAB simulator. To conduct high-fidelity emulation of real-world settings, we use the 802.11n data traces provided by the authors of [30]. The traces contain per-subcarrier (30 subcarriers for 20 MHz) CSI and SNR readings (ranging from 4 to 43 dB) for 18 users. By using the traces, we reconfigure 25 users and a maximum of four AP antennas in our simulations.

Except for the set of fair scheduling protocol simulations, we assume that all users have the same average SNR of 16.7 dB (according to the trace). We set the simulation parameters to the default values in Table 3 and

Table 3 Default simulation parameters

Parameter	Value
Downlink traffic intensity	5 (Mbps)
Uplink traffic intensity	0.5 (Mbps)
K	15
The number of polls used in SUS	10
M	4
SNR	16.7 (dB)
γ	4
o	4

802.11ac [2]. The AP and users generate traffic for each other according to their average sending rates (traffic intensity). We measure the average aggregate throughput of the downlink, uplink^c, and system. Note that the system throughput is defined as the sum of downlink and uplink throughput. We also assume that all protocols used in the simulation enable a packet aggregation scheme.

5.2 802.11ac+ performance

5.2.1 Impact of K

The user diversity will be more effective as the number of users to be considered for the scheduling increases; however, the significant CSI feedback overhead may limit the gain. In order to examine the performance of 802.11ac+, we illustrate the downlink and system throughput of three protocols according to the number of users in Figure 8. Both throughputs increase until the number of users reaches a certain point, after which they start to decrease. 802.11ac+ shows significant throughput gain over the other protocols from the high channel utilization via its user selection scheme. As a result, it achieves the maximum downlink (system) throughput gain of 100% (69%) and 168% (97%) over 802.11ac and SUS, respectively. On

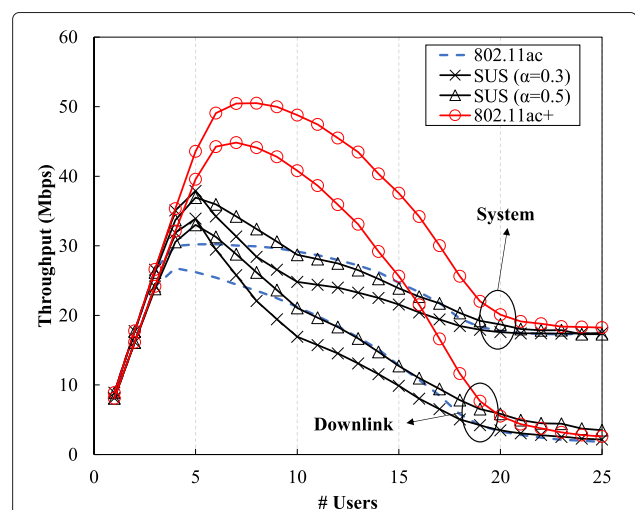


Figure 8 Throughput according to the number of users. Due to dominant uplink traffic and heavy contention, the network throughput decreases with the number of users. We note that α used in SUS controls the trade-off between diversity gain and effective channel gain [6].

the other hand, SUS shows a similar or worse performance compared to 802.11ac except for a moderate number of users. Note that these maximum throughput gains are computed as follows:

$$\max \frac{\text{throughput}_k^{802.11ac+} - \text{throughput}_k^{802.11ac(SUS)}}{\text{throughput}_k^{802.11ac(SUS)}} \quad (18)$$

where $\text{throughput}_k^{\text{scheme}}$ denotes the target scheme throughput when $K = k$.

5.2.2 Impact of M

More AP antennas provide better spatial multiplexing gain while increasing the CSI overhead because a large number of bits are required for representing CSI. In this simulation, we investigate and compare the performance according to the number of AP antennas, as shown in Figure 9. As expected, the throughput of protocols increases with the number of AP antennas. However, the CSI overhead limits the performance of SUS while 802.11ac+ can obtain much higher throughput gain due to small CSI overhead. As a result, it achieves the maximum downlink (system) throughput gain of 98% (51%) and 161% (75%) over 802.11ac and SUS, respectively.

5.2.3 Impact of the number of polls

To investigate the user scheduling gain affected by the CSI overhead, we compare the three system throughput

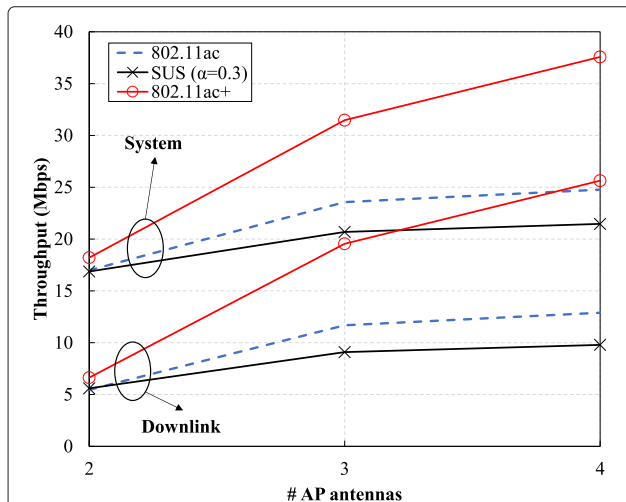


Figure 9 Throughput according to the number of AP antennas. The large number of AP antennas increases the size of the CSI feedback frame. Thus, requesting many users' CSI for scheduling may limit the scheduling gain significantly, as shown in the result of SUS. However, 802.11ac+ can obtain higher throughput gain over both protocols due to the limited feedback.

results according to the number of polls in Figure 10. As expected, SUS suffers from the CSI overhead as the number of polls increases. As a result, SUS shows poor throughput performance compared to 802.11ac after the number of polling frames exceeds seven in the case of $M = 4$. In contrast to SUS, 802.11ac+ outperforms other protocols by using far less CSI feedback while fully harnessing the user scheduling gain. As a result, in the case of $M = 4$ and four polling frames, its throughput gains over 802.11ac and SUS are 51% and 20%, respectively.

5.3 Fair scheduling protocol performance

5.3.1 System throughput

We compare the system throughput of RR-11ac+ and PF-11ac+ with 802.11ac and 802.11ac+ in Figure 11. As shown in the graph, 802.11ac+ and its fair approaches show similar throughputs, although they use different user selection criteria from each other. In particular, the system throughputs of RR-11ac+ and PF-11ac+ are much bigger than that of 802.11ac, since the user scheduling adopted in them still considers the effective channel gain, unlike 802.11ac. Recall that all users experience similar average SNRs in this simulation.

5.3.2 Fairness

To evaluate the fairness performance of RR-11ac+ and PF-11ac+, users are subjected to different average SNRs from 4 to 43 dB, and a user with a higher ID has a bigger SNR than one with a lower ID, unlike in the previous set of simulations.

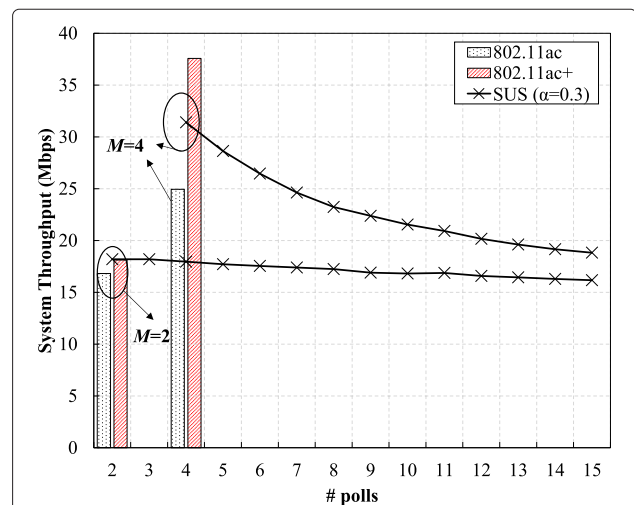


Figure 10 System throughput according to the number of polls. 802.11ac+ obtains higher throughput gain from the user scheduling with a fixed number of polls, while in the case of SUS, the performance degrades as the number of polls increases.

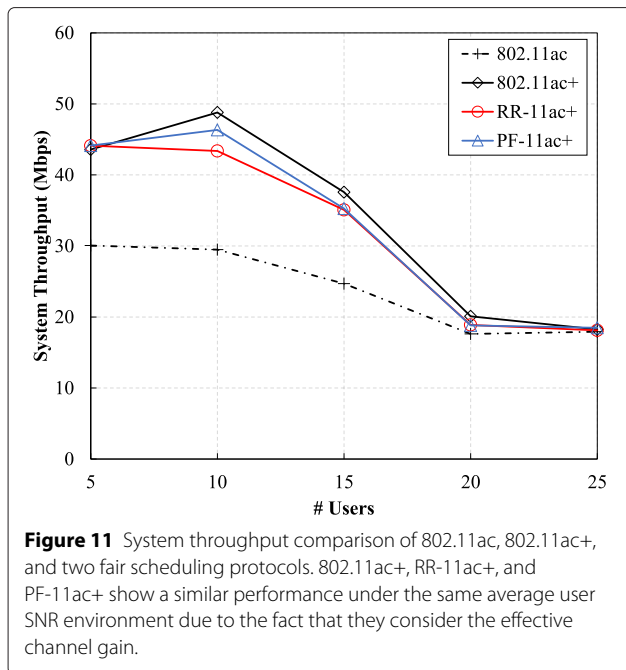
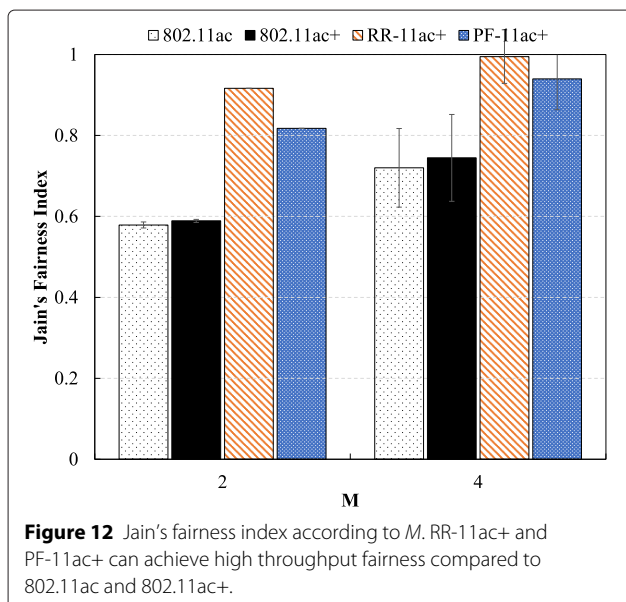


Figure 12 illustrates Jain’s Fairness Index of downlink throughput of four protocols. From the result, we can see that the MU-MIMO transmission with more antennas gives better fairness because it can serve more users at once. Over all cases, RR-11ac+ shows the best fairness performance, followed by PF-11ac+. In particular, RR-11ac+ with $M = 4$ achieves close-to-perfect fairness among users. In the meantime, 802.11ac and 802.11ac+ bring low fairness levels compared with the two fair scheduling protocols. The performance of 802.11ac+ is

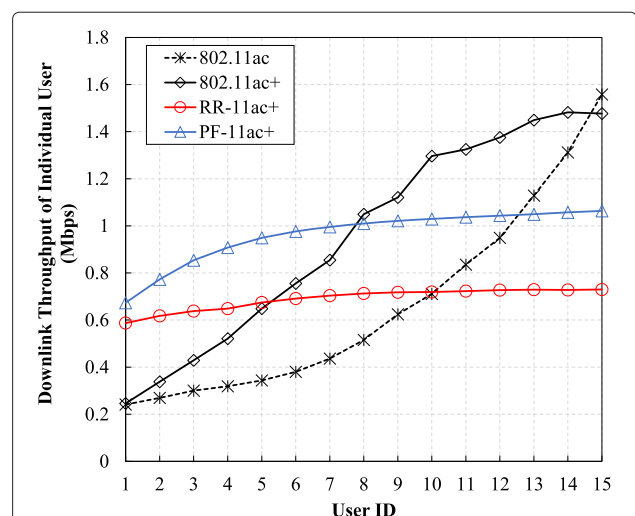


similar to or higher than that of 802.11ac, and yet the sum rate of 802.11ac+ is much higher than that of 802.11ac.

Finally, we plot the downlink throughput that each user attains under each protocol in Figure 13. From the figure, we can clearly see that RR-11ac+ and PF-11ac+ achieve fairly similar throughputs over all users except some users with low SNRs. In addition, in this result, PF-11ac+ shows better aggregate downlink throughput than 802.11ac+, since users in 802.11ac+ have different chances of being scheduled regarding only their ECGs: the user scheduling of 802.11ac+ only favors a user to enhance the sum rate. As a result, there exist some users that have low throughput in 802.11ac+. Surprisingly, in the case of 802.11ac, this phenomenon is more obvious; more users suffer from starvation and a few users with high SNRs enjoy the high throughput. The main reason is that since the user scheduling in 802.11ac+ cares about the channel orthogonality between users, the probability of creating a low-benefit scheduling group is lowered.

6 Conclusions

In this work, we propose a new MU-MIMO MAC protocol, 802.11ac+, which obtains significant user scheduling gain with a far smaller amount of CSI feedback by exploiting channel hint-based polling and active CSI feedback. Trace-driven MATLAB simulation results show that 802.11ac+ achieves downlink throughput gains of up to 100% and 168% over 802.11ac and SUS-based MAC protocols. Additionally, two fair scheduling protocols of 802.11ac+ give close-to-perfect user fairness even when the users experience different channel qualities.



Endnotes

^aThe effective channel vector of the first user is the same as its channel vector, $\mathbf{e}_{(1)} = \mathbf{h}_{(1)}$.

^bIn this paper, we do not consider a per-subcarrier decision. We leave this issue to our future work.

^cFor better graph representation, we omit the graphs for the uplink.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was supported in part by the IT R&D program of MSIP/KEIT. [10041861, Development of Wired/Wireless i-AVB System Technology for Concurrent Transmission of HD Media and Control Data].

Received: 21 October 2014 Accepted: 22 February 2015

Published online: 18 April 2015

References

1. Evolved Universal Terrestrial Radio Access (E-UTRA), Downlink Multiple Input Multiple Output (MIMO) enhancement for LTE-Advanced (Release 11). 3GPP TR V11.0.0. <http://www.3gpp.org/> (2015)
2. IEEE Approved Draft Standard for IT - Telecommunications and Information Exchange Between Systems - LAN/MAN - Specific Requirements - Part 11: Wireless LAN Medium Access Control and Physical Layer Specifications. IEEE P802.11ac/D7.0, 1–456 (2013)
3. G Dimic, ND Sidiropoulos, On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm. *Signal Process. IEEE Trans.* **53**(10), 3857–3868 (2005)
4. J Wang, DJ Love, MD Zoltowski, User selection with zero-forcing beamforming achieves the asymptotically optimal sum rate. *Signal Process. IEEE Trans.* **56**(8), 3713–3726 (2008)
5. M Kobayashi, G Caire, Joint beamforming and scheduling for a multi-antenna downlink with imperfect transmitter channel knowledge. *IEEE J Sel A Commun.* **25**(7), 1468–1477 (2007)
6. T Yoo, A Goldsmith, On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming. *Selected Areas Commun. IEEE J.* **24**(3), 528–541 (2006)
7. S Huang, H Yin, J Wu, VCM Leung, User selection for multi-user MIMO downlink with zero-forcing beamforming. *Vehicular Techn. IEEE Trans.* **62**(7), 3084–3097 (2012)
8. L Jin, X Gu, Z Hu, Low-complexity scheduling strategy for wireless multiuser multiple-input multiple-output downlink system. *Commun. IET.* **5**(7), 990–995 (2011)
9. Z Shen, R Chen, JG Andrews, RW Heath, BL Evans, in *Signals, Systems and Computers, 2005. Conference Record of the Thirty-Ninth Asilomar Conference On*. Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization (Pacific Grove, California, USA, 2006)
10. X Zhang, X Xie, K Sundaresan, in *ACM MOBICOM*. Adaptive feedback compression for MIMO networks (Miami, Florida, USA, 2013)
11. QH Spencer, AL Swindlehurst, M Haardt, Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. *Signal Process. IEEE Trans.* **52**(2), 461–471 (2004)
12. C Hellings, M Joham, W Utschick, in *Smart Antennas (WSA), 2011 International ITG Workshop On*. Gradient-based rate balancing for mimo broadcast channels with linear precoding (IEEE Aachen, Germany, 2011), pp. 1–8
13. C Guthy, W Utschick, in *Proceedings of ITG/IEEE Workshop on Smart Antennas (WSA)*. Sum throughput maximization in quality of service constraint multiuser mimo systems based on perturbation analysis (Vienna, Austria, 2007)
14. FRM Lima, T Ferreira Maciel, WC Freitas, FR Porto Cavalcanti, Improved spectral efficiency with acceptable service provision in multiuser MIMO scenarios. *Vehicular Techn. IEEE Trans.* **63**(6), 2697–2711 (2014)
15. DJ Love, RW Heath, W Santipach, ML Honig, What is the value of limited feedback for MIMO channels? *Commun. Mag. IEEE.* **42**(10), 54–59 (2004)
16. P Frank, A Muller, J Speidel, in *Wireless Conference (EW), 2010 European*. Fair performance comparison between CQI- and CSI-based MU-MIMO, for the LTE downlink (Lucca, Italy, 2010)
17. IEEE Std 802.11n. IEEE Standard for Information technology - Local and metropolitan area networks - Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE P802.11 (2009)
18. K Huang, RW Heath, JG Andrews, Limited feedback beamforming over temporally-correlated channels. *Signal Process. IEEE Trans.* **57**(5), 1959–1975 (2009)
19. V Pohl, PH Nguyen, V Jungnickel, C Von Helmolt, in *IEEE Global Telecommunications Conference*. How often channel estimation is needed in MIMO systems (San Francisco, USA, 2003)
20. W Wang, A Harada, Enhanced limited feedback schemes for DL MU-MIMO ZF precoding. *Wireless Commun. IEEE Trans.* **12**(4), 1554–1561 (2013)
21. JH Lee, W Choi, Optimal feedback rate sharing strategy in zero-forcing MIMO broadcast channels. *Wireless Commun. IEEE Trans.* **12**(6), 3000–3011 (2013)
22. DJ Love, RW Heath, VKN Lau, D Gesbert, BD Rao, M Andrews, An overview of limited feedback in wireless communication systems. *Selected Areas Commun. IEEE J.* **26**(8), 1341–1365 (2008)
23. EJ Candes, MB Wakin, An introduction to compressive sampling. *Signal Process. Mag. IEEE.* **25**(2), 21–30 (2008)
24. ST Qaseem, TY Al-Naffouri, S Alghadhban, in *Telecommunications (ICT), IEEE International Conference On*. Compressive sensing for feedback reduction in MIMO broadcast channels (Doha, Qatar, 2010)
25. D Gesbert, M-S Alouini, in *Signal Processing and Information Technology, IEEE International Symposium On*. Selective multi-user diversity (Darmstadt, Germany, 2003)
26. D Gesbert, M-S Alouini, in *Communications, IEEE International Conference On*. How much feedback is multi-user diversity really worth? (Paris, France, 2004)
27. V Hassel, D Gesbert, M-S Alouini, GE Oien, A threshold-based channel state feedback algorithm for modern cellular systems. *Wireless Commun. IEEE Trans.* **6**(7), 2422–2426 (2007)
28. W Zhang, KB Letaief, MIMO broadcast scheduling with limited feedback. *Selected Areas Commun. IEEE J.* **25**(7), 1457–1467 (2007)
29. Q Yang, X Li, H Yao, J Fang, K Tan, W Hu, J Zhang, Y Zhang, in *Proc. of ACM SIGCOMM*. Bigstation: enabling scalable real-time signal processing in large MU-MIMO systems (Hong Kong, China, 2013)
30. A Bhartia, Y-C Chen, S Rallapalli, L Qiu, in *ACM MOBICOM*. Harnessing frequency diversity in Wi-Fi networks (Las Vegas, Nevada, USA, 2011)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com