This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2018.2879668, IEEE Wireless Communications Letters

Practical User Selection with Heterogeneous Bandwidth and Antennas for MU-MIMO WLANs

Sulei Wang, Zhe Chen, Yuedong Xu, Xin Wang, and Qingsheng Kong

***Abstract*—User selection is one of the most important components for next generation multi-user multiple-input-multiple-output (MU-MIMO) WLANs. However, state-of-the-art approaches neglect the heterogeneity of users in the available bandwidth and the number of antennas, which diminishes their performance considerably. To tackle this challenge, we formulate a novel integer optimization framework to select the antennas of heterogeneous users simultaneously. With estimated signal-to-interference-and-noise ratio (SINR) of users via channel vector projection, we propose a low-complexity branch-and-prune algorithm to search for the near-optimal combinations of user antennas. Our algorithm is compatible with legacy 802.11ac and is implemented on the software defined radio system. Extensive experiments show that our algorithm achieves around 95% of the optimal throughput and outperforms a benchmark scheme with a 1.18× gain in realistic indoor environments.**

***Index Terms*—MU-MIMO, User Selection, Heterogeneity, Branch-and-prune.**

I. Introduction

MU-MIMO is a key enabling technology to scale up the capacity of 802.11ac wireless local area networks (WLANs) [1]. Equipped with multiple antennas, an access point (AP) is capable of transmitting multiple data streams to different users or receive antennas concurrently, achieving a spatial reuse gain up to the number of transmit antennas [2]. In practice, this spatial reuse gain depends on the channel orthogonality among users and such orthogonality cannot be always preserved [3]. Hence, the AP needs to schedule the transmission of a group of users wisely in each slot so as to reduce the inter-user interference, especially when the candidate pool is large.

The design of user selection strategy has gripped much attention recently in 802.11ac WLANs. Authors in [4] proposed an orthogonality probing based user selection scheme named *OPUS*. *SIEVE* [5] balanced the trade-off between performance and complexity with a scalable multiuser selection module. Recently, *MUSE* [6] was designed to perform user selection with limited CSI feedback on commodity Wi-Fi devices.

Although the above seminal approaches perform well, they neglect a crucial phenomenon, that is, *user heterogeneity*. The origin of user heterogeneity stems from the diversity of devices accessing WLANs, e.g. laptops, cellphones, wearable as well as smart-home devices. The users possess different number of antennas and support different maximum bandwidths. The user heterogeneity poses new challenges to the user selection in MIMO systems. For instance, grouping all the antennas of one multiple-antenna user does not always yield a high throughput. If two users are selected whose bandwidths are 20MHz and 40MHz respectively, the AP will transmit in the 20MHz bandwidth. Therefore, the user selection strategy must be capable of handling the user heterogeneity as it is genuinely remarkable and coupled with the mitigation of inter-user interference.

In this paper, we formulate the user selection as an integer programming problem. The objective is to maximize the aggregate throughput in each slot, constrained by the number of transmit antennas, and the available bandwidth of users. Finding the optimal set of users incurs a prohibitive computational complexity, thus is infeasible for online processing. To circumvent this difficulty, we first adopt a channel vector projection method [7] to estimate the SINR of user antennas. A branch-and-prune algorithm [8] is further applied to select the user antennas incrementally. Specifically, our algorithm maintains multiple candidate user antenna combinations that enlarge the search space for better throughput. The algorithm is compatible with legacy 802.11ac and advanced techniques including CSI compression. We implement the proposed algorithm on the software defined radio platform WARP [9], and extensive experiments manifest the effectiveness.

II. Problem Formulation

Suppose that the AP is equipped with S antennas and there are K users contending for transmission, where user k is equipped with n_k antennas and supports the maximum available bandwidth B_k. With the binary selection mask $x_{s,k,i}$ indicating whether antenna i at user k is selected and served by data stream s or not, we denote the receiving SINR as $SINR_{s,k,i}$. The corresponding throughput $U_{s,k,i}$ is given by $U_{s,k,i} = B\log(1 + SINR_{s,k,i})$ where B is the transmission bandwidth. Thus the problem is formulated as follows:

$$\max_{\mathbf{x},B} \quad \sum_{s=1}^{S}\sum_{k=1}^{K}\sum_{i=1}^{n_k} x_{s,k,i}U_{s,k,i} \tag{1}$$

$$\text{s.t} \quad \sum_{s=1}^{S}\sum_{k=1}^{K}\sum_{i=1}^{n_k} x_{s,k,i} \le S, \tag{2}$$

$$B = \min \bigcup_{s,k,i}\{x_{s,k,i}B_k\}\backslash\{0\}, \tag{3}$$

$$x_{s,k,i} \in \{0,1\}. \tag{4}$$

Eq. (2) means that the S-antenna AP can serve no more than S user antennas concurrently. The users support diverse maxi-

This work is supported by Natural Science Foundation of China (No. 61772139), Shanghai-Hong Kong Collaborative project (No. 18510760900) and CERNET Innovation Project (No. NGII20170209). *Yuedong Xu is the corresponding author.*

Sulei Wang, Yuedong Xu and Qingsheng Kong are with Research Center of Smart Networks and Systems, School of Information Science and Technology, Fudan University, Shanghai, China (email: {wangsl16, ydxu, qskong}@fudan.edu.cn).

Zhe Chen and Xin Wang are with School of Computer Science, Fudan University, Shanghai, China (email: {zhechen13, xinw}@fudan.edu.cn).

mum bandwidths, while the AP can only transmit on a central frequency with one channel bandwidth in each transmission slot. Therefore, once the set of user antennas are selected, the AP transmits using the lowest available bandwidth that can be supported by them (Eq. (3)). For example, if three user antennas supporting up to 20MHz, 40MHz, and 40MHz are selected, the actual bandwidth for transmission is 20MHz.

Denote N by the total number of user antennas, i.e. $N = \sum_{k=1}^{K} n_k$. There exist $\sum_{T=1}^{S} \binom{N}{T}$ possible user antenna combinations. Exhaustively searching for the optimal combination entails a prohibitively high complexity, which is not suitable for online computation in MIMO WLANs.

III. Proposed Algorithm

In this section, we first introduce a lightweight method to estimate the SINR of user antennas. Then we adopt a low-complexity branch-and-prune algorithm to select user antennas incrementally. The complexity of our algorithm is analyzed and the feasibility of practical implementation is described.

A. SINR Inference

The prerequisite of designing a user selection algorithm is to infer the SINR of possible user antenna combinations. Here, we use an example in Fig. 1 to illustrate how the channel vector projection can achieve this goal. The two-antenna AP transmits signals in a two-dimensional space where $h_{11}, h_{12}, h_{21}, h_{22}$ are complex channel gains between AP and the user antennas. When precoding a symbol for U_2, the AP nullifies the interference from U_1 by projecting the symbol for U_2 in a direction orthogonal to $\mathbf{h_1}$. This projection process leads to a degradation in SINR of U2, which is reflected on the reduction of the length of blue vector. The SINR reduction decreases as the inter-user reception angle θ increases. As a special case $\theta = 90°$, the channel vectors of U_1 and U_2 are perfectly orthogonal to each other, hence there is no SINR reduction.

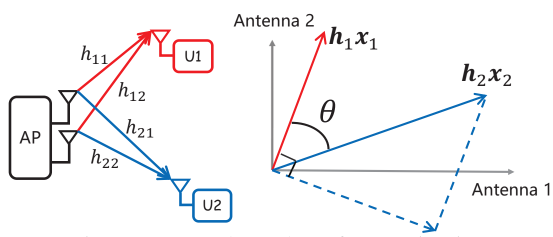

Fig. 1: SINR degrades after projection

Denote $SINR_{orig}$ by the SINR of a user antenna when it is served alone, then the SINR after projection is given by

$$SINR_{proj} = \sin^2(\theta) \cdot SINR_{orig}. \tag{5}$$

Generally, if an AP equipped with S antennas serves m user antennas ($m < S$) and a new antenna is added into the group, the SINR reduction of the new antenna can also be obtained by Eq. (5). In this case, $\sin\theta$ is obtained by:

$$\sin\theta = \frac{|\mathbf{h}_\perp \cdot \mathbf{h}|}{||\mathbf{h}_\perp|| \cdot ||\mathbf{h}||} \tag{6}$$

where $\mathbf{h}$ is the channel vector of the new user and $\mathbf{h}_\perp$ is the vector orthogonal to the subspace spanned by the channel vectors of the m selected users.

B. Greedy Search with Pruning

A conventional wisdom of greedy algorithm is to select a user from the candidate pool in each step that yields the highest throughput increment. However, there is no remedy when the first several steps deviate far away from optimality. In our problem, a simple greedy approach may lead to the improper downgrading of bandwidth or assignment of antennas. We employ the branch-and-prune algorithm [8] to perform user selection for MU-MIMO WLANs. We keep M candidate user antenna combinations instead of selecting only one and rejecting all the others in each incremental step, which distinguishes our algorithm from previous methods.

The high-level operation flow of our algorithm is as follows:

1) **Initialization.** Each of N user antennas is initialized as a candidate user antenna combination, generating N candidates.
2) **Branching.** For each candidate user antenna combination, we add all the user antennas that have not been included in it respectively.
3) **Pruning.** We compute the sum throughput of all the branched user antenna combinations, keep the top-M user antenna combinations and reject all the others.
4) **Looping.** The above branching and pruning process repeats until all the antennas of the AP have been assigned. We then select the user antenna combination with the maximum sum throughput in the final M choices.

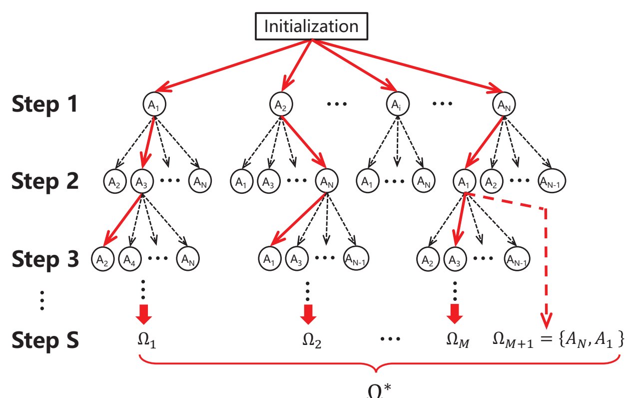

Fig. 2: Greedy user selection with pruning

Fig. 2 provides an illustrative example of our algorithm. A_i ($i = 1, \cdots, N$) refers to the i^{th} antenna. In this tree, each node represents a candidate user antenna combination. The arrows indicate the incremental user selection process. An arrow pointing to a node means that an additional user antenna is added. The red line arrow means that the node is branched and retained while the black dashed arrows point to the pruned nodes. The number of red line arrows in each step M is a tuneable parameter in which the branch-and-prune algorithm is a naive greedy approach at $M = 1$. Here, we let M be 3 so that only three combinations are retained in each step.

In the initialization step, each user antenna is initialized as a candidate combination itself. Later on, each step contains a branching phase and a pruning phase. In the branching phase, we add each remaining user antenna in the set of candidate antennas. For instance, when each node A_i is added to $\{A_1\}$ for $i \neq 1$, we obtain $N - 1$ branches of user antenna combinations, i.e. $\{A_1, A_2\}$ $\cdots$, $\{A_1, A_N\}$. Hence, there are

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2018.2879668, IEEE Wireless Communications Letters

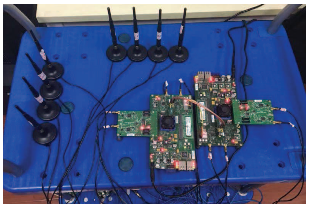

Fig. 3: WARP platform

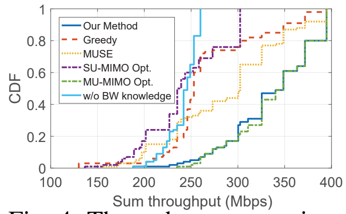

Fig. 4: Throughput comparison

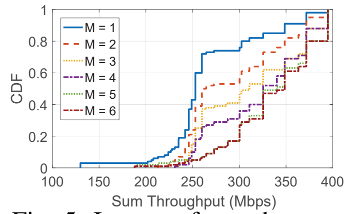

Fig. 5: Impact of search space

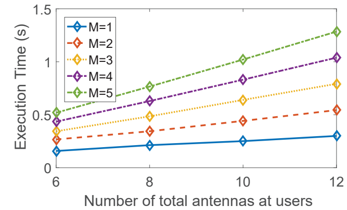

Fig. 6: Execution time

$N(N-1)$ branches in total at the first step. In the pruning phase, the total throughput of each user antenna combination is computed. The M user antenna combinations with highest throughput are retained, i.e. $\{A_1, A_3\}$, $\{A_2, A_N\}$ and $\{A_N, A_1\}$, and the remaining branches are pruned subsequently.

The branch-and-prune process terminates automatically when the number of selected user antennas equals to the number of antennas at the AP. Since only one user antenna is added in each incremental step, the algorithm will terminate at Step S. We then select the user antenna combination yielding the maximum sum throughput in the final M choices Ω_i $(i = 1, \cdots, M)$.

Note that a combination with more user antennas does not always lead to a higher total throughput. There are two reasons accounting for this phenomenon: 1) the channel of a newly added user antenna is highly correlated with the existing user antennas; 2) the newly added user antenna that is of lower bandwidth may force the existing users to use the lower bandwidth, causing a downgraded throughput. When a candidate user antenna combination outperforms all its branching combinations, we mark it as a special combination and let it compare with the top-M combinations in the final step. The branch-and-prune loop remains unchanged. As shown in Fig. 2, $\Omega_{M+1} = \{A_N, A_1\}$ is such an user antenna combination, and is put in the final candidate pool.

Fairness Control: To maintain fairness among user antennas in each scheduling round, we divide all the user antennas into an active set and an inactive set. The user antennas in the active set are moved to the inactive set after being selected. The user antennas in the inactive set are restored to the active set when it is empty. This mechanism is easy to be implemented and can effectively ensure the fairness among different user antennas.

Complexity: Our user antenna selection algorithm has a polynomial-time computational complexity. In the initialization step, N branches are created and maintained without any pruning. Therefore, $N(N-1)$ candidate combinations are branched at step 1. The selection of the highest M combinations yields a complexity order $O(MN^2)$. After step 1, only M branches are maintained until the end of the branching of all the steps. Hence, there exist $M(N-1)$ branches in each following step where the selection of the top-M combinations has a complexity order $O(M^2N)$. Considering that the above procedure is executed for $S-1$ times since Step 2, our branch-and-prune algorithm has a complexity order of $O(MN^2 + SNM^2)$, which is lower than exhaustive searching.

Compatibility: Our algorithm does not require any modification on vanilla 802.11ac medium access control (MAC) protocols. The algorithm is executed based on the CSI feedback, and the data streams are precoded based on the selection results. It is compatible with techniques including CSI feedback compression and frame aggregation, as long as effective CSI feedback is provided.

IV. Implementation and Evaluation

A. Implementation and Experimental Setup

We implement our user selection algorithm on the software defined radio platform WARP [9]. The PHY layer follows the 802.11ac specifications, consisting of OFDM, modulation/coding, channel estimation and ZFBF. In the MAC layer, we implement the aforementioned CSI feedback mechanism. Our experiments are conducted in a typical office environment. The users are placed randomly and sometimes move at a walking speed for evaluating the mobility scenario.

B. Evaluation

1) Effectiveness of our algorithm: To evaluate the effectiveness of our algorithm, we setup a MU-MIMO WLAN with a four-antenna AP and a random number of heterogeneous users that possess twelve antennas in total. Based on the SINR calculated from the received preambles, we obtain the system throughput via the SNR-to-rate lookup table [2]. For comparison, we implement a greedy user selection algorithm $(M = 1)$. Besides, we implement *MUSE* [6], a heuristic user grouping scheme for MU-MIMO with the consideration of bandwidth heterogeneity. Moreover, we calculate the optimal user selection results in SU-MIMO and MU-MIMO modes offline. Our experiments are conducted in one hundred different scenarios, and each experiment has a runtime of 100 rounds.

Fig. 4 plots the CDFs of the sum throughput. The optimal result in SU-MIMO mode is the worst, verifying the necessity of MU-MIMO. The greedy algorithm is also not satisfactory, because the AP has no knowledge of the inter-user interference of unselected users when performing user selection incrementally. Our user selection algorithm achieves a total median throughput of 336.675 Mbps, outperforming *MUSE* by 18%. Less than 20% of experiments have a throughput below 300 Mbps in our algorithm, while more than 40% of them have in MUSE. The proposed algorithm is also repeated without considering bandwidth heterogeneity. The resulting throughput is much worse than that with bandwidth knowledge, implying the significance of taking bandwidth heterogeneity into consideration. Furthermore, our algorithm reaches around 95% of throughput of the optimal result in MU-MIMO mode.

2) Impact of search space: We hereby evaluate the proposed branch-and-prune algorithm with different scales of search space. Recap that the search space is determined by M, the number of the candidate user combinations. Fig. 5 illustrates the CDF of sum throughput when M increases from one to six. A larger M yields a better sum throughput at the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2018.2879668, IEEE Wireless Communications Letters

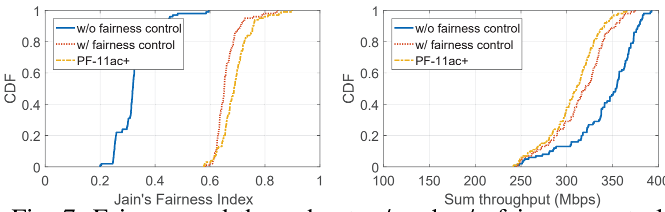

Fig. 7: Fairness and throughput w/ and w/o fairness control

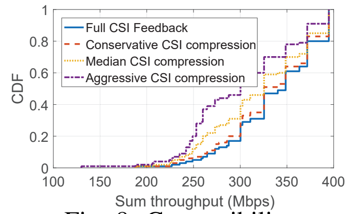

Fig. 8: Compatibility

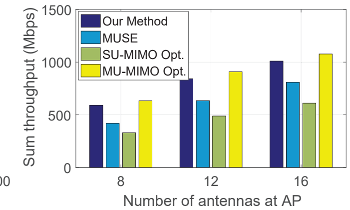

Fig. 9: Large-scale simulation

cost of increased computational complexity. When M changes from one to five, one can witness a nearly 20% throughput gain in most of the experiments; when it increases from five to six, the throughput gain becomes negligible. Hence, an optimistic message on our algorithm is that a small search space (e.g. M is chosen to be five) might be good enough.

3) Execution time: We also record the execution time under different numbers of user antennas and different search space. As shown in Fig. 6, the execution time grows almost linearly with the number of user antennas, but fortunately with gentle slopes. Meanwhile, the execution time is proportional to the search space. Compared with the greedy algorithm ($M = 1$), the proposed algorithm costs more runtime, but leads to better throughput performance. Note that this part of experiments is conducted on a laptop configured with Intel Core i7-4500U 1.80GHz CPU using MATLAB. We believe that a binary code implementation on commercial AP will be much faster.

4) Effectiveness of fairness control: We further test our user and antenna selection algorithm with fairness consideration. The preceding experiments are repeated except that the fairness control scheme is activated. For comparison, we implement the proportional fairness scheduling of *PF-11ac+* in [10]. We adopt Jain's fairness index to quantify the fairness of all the receive antennas where the CDFs of average throughput and fairness index are shown in Fig. 7. One can clearly observe that our fairness control method balances the transmission opportunities well among all the receive antennas. *PF-11ac+* achieves better fairness performance because it is designed for long-term fairness. However, enforcing fairness control usually reduces the sum throughput. Both our fairness control method and PF-11ac+ degrades the sum throughput slightly, because some users with poor channel conditions or low bandwidth are grouped for fairness guarantee.

5) Compatibility: The proposed algorithm also works under CSI compression mechanisms that intends to reduce the feedback overhead. We implement *AFC*, a representative CSI compression mechanism proposed in [11] and fix the configurations under three different compression levels for verifiable comparison: conservative CSI compression (sharing CSI across 10ms, one subcarrier and quantizing numerical values into 8 bits), median CSI compression (20 ms/two subcarriers/6 bits) and aggressive CSI compression (40 ms/four subcarriers/4 bits). Under the same experimental setup as before (four-antenna AP and twelve receive antennas), *AFC* reduces the overhead from 3.364ms to 0.589ms, 0.131ms and 0.029ms under different compression levels. However, we can see from Fig. 8 that the conservative compression has almost negligible impact on system throughput, while the median compression and the aggressive compression experience 7.3% and 15.4% throughput loss respectively.

C. Large-scale Simulation

To evaluate the scalability of our algorithm, we consider a large scale WLAN that consists of an AP with eight to sixteen antennas and a certain amount of users with a total number of fifty receive antennas. Due to the restriction of SDR platform (expensive price and up to four antennas on each board), we collect realistic wireless transmit and receive traces using our system at different locations, and then emulate the selection procedure offline. Fig. 9 shows the mean throughputs of MUSE, the optimal scheme and our algorithm. Our algorithm outperforms optimal SU-MIMO and MUSE by 78.7% and 40.9%, respectively. The performance gap between our algorithm and the optimal MU-MIMO is nearly 10%, slightly larger than that in Fig. 4. When the AP is equipped with more antennas, e.g. sixteen, our algorithm still exhibits remarkable gains and remains close to the optimality. This manifests that our algorithm possesses an excellent scalability.

V. Conclusion

In this work, we address the user selection problem in MU-MIMO WLANs with heterogeneous maximum bandwidth and number of receive antennas. A novel branch-and-prune algorithm is proposed to achieve the low-complexity selection of user antennas. Our algorithm is compatible with legacy 802.11ac, and is implement on the software defined radio platform WARP. Experimental results demonstrate that our algorithm outperforms a very recent counterpart by 1.18 times and reaches around 95% of the optimal results.

References

[1] "IEEE 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 - Amd 4: Enhancements for Very High Throughput for operation in bands below 6GHz," IEEE P802.11ac/D3.0, June 2012.
[2] Matthew Gast. 2013. 802.11ac: A Survival Guide (1st ed.). O'Reilly Media, Inc..
[3] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-Antenna Downlink Channels with Limited Feedback and User Selection," IEEE Journal on Selected Areas in Communications, vol. 25, no. 7, 2007.
[4] X. Xie and X. Zhang, "Scalable User Selection for MU-MIMO Networks," in Proc. of IEEE INFOCOM, 2015.
[5] W. Shen, K. Lin, M. Shen and K. Tan, "SIEVE: Scalable User Grouping for Large MU-MIMO Systems," in Proc. of IEEE INFOCOM, 2015.
[6] S. Sur, I. Pefkianakis, X. Zhang and K. Kim, "Practical MU-MIMO User Selection on 802.11ac Commodity Networks," in Proc. of ACM MobiCom, 2016.
[7] W. Shen, Y. Tung, K. Lee, K. Lin, S. Gollakota, D. Katabi and M.Chen, "Rate Adaptation for 802.11 Multiuser MIMO Network," in Proc. of ACM MobiCom, 2012.
[8] J. Porta, L. Ros, F. Thomas, and C. Torras, "A branch-and-prune algorithm for solving systems of distance constraints," in Proc. of IEEE ICRA, 2003.
[9] WARP Project, http://warpproject.org
[10] K. Lee and C. Kim, "User scheduling for MU-MIMO transmission with active CSI feedback," in EURASIP Journal on Wireless Communications and Networking, 2015.
[11] X. Xie, X. Zhang, and K. Sundaresan, "Adaptive Feedback Compression for MIMO Networks," in Proc. of ACM MobiCom, 2013.