Normal Graphs for Downlink Multiuser MIMO Scheduling

Jung-Chieh Chen†*, Cheng-Hsuan Wu††, Chao-Kai Wen‡, Yao-Nan Lee†††, Hsin-Yi Lu‡, and Pangan Ting‡‡
†Dept. of Optoelectronics & Comm. Eng., National Kaohsiung Normal University, Kaohsiung 802, Taiwan.
††Institute of Communications Engineering, National Tsing Hua University, Hsinchu 300, Taiwan.
‡‡Industrial Technology Research Institute, Hsinchu 310, Taiwan
†††MediaTek Inc., Hsinchu 300, Taiwan
*E-mail: jcchen@nknucc.nknu.edu.tw

Abstract—Inspired by the success of the low-density parity-check (LDPC) codes in the field of error-control coding, in this paper we propose transforming the downlink multiuser multiple-input multiple-output scheduling problem into an LDPC-like problem using the normal graph. Based on the normal graph framework, soft information, which indicates the probability that each user will be scheduled to transmit packets at the access point through a specified angle-frequency sub-channel, is exchanged among the local processors to iteratively optimize the multiuser transmission schedule. Computer simulations show that the proposed algorithm can efficiently schedule simultaneous multiuser transmission which then increases the overall channel utilization and reduces the average packet delay.

Keywords: Normal Graph, Multiuser MIMO Scheduling

I. INTRODUCTION

Downlink multiuser multiple-input multiple-output (MU-MIMO) systems have attracted considerable attention recently for their potential to increase the system capacity. However, in a multiuser scenario, a limited amount of system resources is shared among all the users within a service area. Recent research results, e.g., dirty-paper coding scheme (DPCS) in [1], have shown that a high system throughput (sum rate) is achievable in an MU-MIMO system if both the transmitters and the receivers have perfect knowledge of the channel state information (CSI). From a practical system implementation perspective, it is almost impossible for an access point (AP) to acquire the exact CSI of all users through feedback channels. Therefore, instead of adopting an unrealizable DPCS in a downlink MU-MIMO system, we propose utilizing a low-complexity angle-frequency coding scheme (AFCS) [2] for simultaneous multiuser transmission.

When a group of subscribers simultaneously want to receive packets from an AP, the mechanisms in the medium access control (MAC) layer also play an important role to prevent severe multiuser interference at the subscriber sides. Therefore, in this paper, using the AFCS system resource partitioning technique [2], we present a design with a low-complexity algorithm for scheduling simultaneous multiuser transmissions in order to 1) prevent severe multi-user interference at the subscriber sides and 2) efficiently utilize system resources in a downlink MU-MIMO wireless system. The proposed scheduling algorithm is based on the modeling and computational methodology of normal graphs [3]. In a normal graph framework, the soft information, which indicates the probability that each user will be scheduled to transmit packets through a specified angle-frequency sub-channel, is exchanged among the local processors to iteratively optimize the downlink multiuser transmission schedule. From computer simulations, the proposed normal graph downlink scheduling algorithm is shown to efficiently schedule simultaneous multiuser transmission which then increases the overall channel utilization and reduces the average packet delay.

II. BACKGROUND AND PROBLEM DEFINITION

In this work, we call the MU-MIMO system with the AFCS proposed in [2] the MU-AFCS system. By “system resources,” we mean the available angle-frequency sub-channels at the AP side. The detailed operation of the MU-AFCS system is shown in Fig. 1. At the AP side, pilot signals are transmitted through each of the $M_T$ Fourier-basis beamforming vectors [2] at each frequency sub-carrier, where $M_T$ is the number of transmission antennas, there being a total of $N$ frequency sub-carriers. At the subscriber sides, each subscriber should measure the received power of the pilot signals transmitted from the $M_T$ Fourier-basis beamforming vectors at each frequency sub-carrier. Therefore, a 2-D CSI look-up-table (LUT) (or 2-D received signal strength pattern) for each subscriber can be constructed, as shown in Fig. 1.

![System diagram of the MU-MIMO wireless communication system.](image)

In general, the $(i,j)^{th}$ entry of the 2-D CSI LUT contains the channel response of the $j^{th}$ transmission beam and the $j^{th}$ frequency sub-carrier, where the horizontal entry represents the channel response in the frequency domain while the vertical
entry represents the channel response in the angle (beam) domain. The channel response in the 2-D CSI LUT indicates the power level that the subscriber can receive if the signals are transmitted through the corresponding angle-frequency sub-channel at the AP side. In order to reduce the complexity of the 2-D CSI LUT, we can use the binary values “1” and “0” to represent the link quality at the corresponding angle-frequency sub-channel. Each subscriber can construct a binary-valued 2-D CSI LUT by estimating the signal-to-interference-and-noise-ratio (SINR) at all the angle-frequency sub-channels with the help of the pilot signals. If the SINR ratio of an angle-frequency sub-channel exceeds a pre-determined threshold for establishing a reliable link, the entry of the CSI LUT corresponding to this angle-frequency sub-channel will be assigned a “1”; otherwise, a “0” will be assigned. For example, with a subscriber \( k \), an \( M_T \times N \) binary-valued CSI LUT \( F^k \) should be constructed, where \( F^k = \begin{bmatrix} f_{1}^k, f_{2}^k, \ldots, f_{N}^k \end{bmatrix} \), with \( f_{n}^k \) denoting the status of the angle sub-channel in the \( n \)th frequency sub-carrier. Thus, the entry of the 2-D CSI LUT can be represented by

\[
f_{n,m}^k = \begin{cases} 
1, & \text{if this angle-frequency sub-channel can establish a reliable link for subscriber } k \\
0, & \text{otherwise}
\end{cases}
\]  

At the AP side, when the AP receives all the subscribers’ binary-valued CSI LUTs, we should design an efficient algorithm in the MAC layer to efficiently schedule all the simultaneous multi-user transmissions without severe multi-user interference at the subscriber sides. We should also exploit all the angle-frequency resources.

The concerns for avoiding severe multi-user interference from the binary-valued CSI LUT are as follows. First, if we can schedule the subscribers’ simultaneous transmissions in different “1” pattern of \( \{ f_{n}^k \} \) at the \( n \)th frequency sub-carrier, the packets of selected users can then be transmitted simultaneously without severe multi-user interference at the \( n \)th frequency sub-carrier. The reason is that if two users have an overlapping pattern in “1” and indeed transmit signals through these “1” pattern sub-channels, the transmission signal of one user will be “heard” by another user. These undesired signals result in severe multi-user interference. Unlike some other works [1], where perfect CSI is required, our 2-D binary-valued CSI LUT enjoys not only a low feedback rate but also low complexity.

After the scheduler finishes the scheduling task, a transmission scheduling pattern can be obtained. The transmission scheduling pattern shows which subscribers are allowed to transmit through which angle-frequency sub-channel in a time slot. Without loss of generality, we assume that a time slot equals the required time for signals to transmit over the sub-channel.

For ease of exposition, let us represent the scheduling pattern using an \( M_T \times N \) binary matrix \( S \). The \((m,n)\)th entry of matrix \( S \) can then be expressed as

\[
s_{m,n} = \begin{cases} 
1, & \text{if (m,n)} \text{th angle-frequency sub-channel is used by some subscriber} \\
0, & \text{otherwise}
\end{cases}
\]  

Next, before we give the definition of a throughput (or channel utilization), the following assumptions are made: We decompose the transmission data burst into packets before sending them from the AP to subscribers. A packet is defined as a single unit of data that can be sent through one angle-frequency sub-channel. The length of a time slot equals the required time to transmit a packet. Therefore, in each angle-frequency sub-channel, only one packet is allowed to be transmitted in a time slot. Based on the above assumptions, the channel utilization for the entire system per time slot is given by

\[
\rho = \frac{\sum_{n=1}^{N} \sum_{m=1}^{M_T} s_{m,n}}{M_T \times N}.
\]  

III. THE ALGORITHM

A. Normal Graph Modeling

A normal graph [3], consisting of nodes and edges, is a graphic representation of a group of mutually-interactive check rules. Nodes and edges correspond to local rules and variables, respectively. As long as we can transform a problem into a normal graph and specify all the local rules enforced by all the nodes, the problem can then be easily solved using a standard procedure, the soft-information-passing sum-product algorithm. Normal graph modeling for the downlink MU-MIMO scheduling problem is realized through the following two steps:

1. We define two types of nodes: a square node and a circular node. The square represents a subscriber and the circular node represents an angle sub-channel. We will take the CSI LUT shown in Fig. 2 as an example. Nodes \( SS_1, SS_2 \), and \( SS_3 \) represent three subscriber nodes, and nodes \( AS_1, AS_2, AS_3, \) and \( AS_4 \) represent four angle sub-channel nodes.

2. In Fig. 2, an edge is connected from a subscriber node to an angle sub-channel node whenever the user can transmit packets through that angle sub-channel according to the CSI LUT. These edges altogether represent a codeword. Each edge is associated with a codeword bit, which can be either a “1” or “0”. The bit “1” indicates that the subscriber is scheduled to transmit a packet through the angle sub-channel; “0” indicates that the subscriber will not be scheduled to transmit a packet through the angle sub-channel. We define \( A(i) \) as the set of subscriber nodes that are connected to angle sub-channel \( i \) and define \( S(j) \) as the set of angle sub-channel nodes that are connected to subscriber \( j \).

In other words, each codeword represents the assignment of the angle sub-channel in a frequency sub-carrier. We just
need to determine the best codeword according to the local constraints. In order to avoid severe multiuser interference at the subscriber sides, we have to impose local constraints at each node, as will be explained next.

B. Local Constraints

There are two legal codewords for each subscriber node: One is an all-zero codeword and another is all-one codeword. An all-one codeword indicates that a subscriber will be scheduled to transmit a packet through the associated angle sub-channels; otherwise, the subscriber will not be scheduled to transmit a packet at this frequency sub-channel. We will take five angle sub-channels connected to a subscriber as an example. The legal codewords for this subscriber are either (1,1,1,1,1) or (0,0,0,0,0). For ease of exposition, we define the codebook that stores the legal codewords (or the local constraints) for subscriber \( j \) to be \( C(j) \).

![Diagram](image)

Fig.2 The normal graph model for the downlink MU-MIMO scheduling problem.

C. Soft Information Calculation

The soft information at each node can be calculated as follows:

- **Initialization:** The a priori probability \( P \), i.e., the initial SI that \( AS \) \( i \) passes to \( SS \) \( j \) as the probability that \( AS \) \( i \) will be assigned by \( SS \) \( j \)

\[
\text{SI}(AS \ i, SS \ j, 1) = 1 - \text{SI}(AS \ i, SS \ j, 0) = P ,
\]

where \( \text{SI}(x,y,s) \) denotes the soft information passed from node \( x \) to node \( y \), with the codeword bit associated with the edge connecting the two nodes being \( s \), and \( P \) is uniformly and randomly distributed over the interval \((0,1)\).

- **From SS to AS:** After each SS receives SIs from its associated ASs, it calculates its output SIs based on its codebook and passes the SIs back to the associated ASs. According to the sum-product algorithm, the SI passed from \( SS \) \( j \) to \( AS \) \( i \) can be expressed as

\[
\text{SI}(SS \ j, AS \ i, s) = \lambda_y \left[ 1 - \sum_{l \in C(j)} \text{SI}(AS \ l, SS \ j, s) \right] \prod_{l \in C(j)} \text{SI}(AS \ l, SS \ j, s)
\]

where \( \lambda_y \) is a constant to make sure that \( \text{SI}(AS \ i, SS \ j, 0) + \text{SI}(AS \ i, SS \ j, 1) = 1 \).

- **From AS to SS:** In order to complete an iteration loop, each AS needs to calculate its output SIs with the output SIs from its associated SSs. Therefore, the SI passed from \( AS \) \( i \) to \( SS \) \( j \) is given by

\[
\text{SI}(AS \ i, SS \ j, 1) = \lambda_y \prod_{l \in A(i,j)} \text{SI}(SS \ l, AS \ i, 0),
\]

\[
\text{SI}(AS \ i, SS \ j, 0) = \lambda_y \sum_{k=0}^{l \in A(i,j)} \text{SI}(SS \ j, AS \ k, 1) \times \prod_{l \in A(i,j)/k} \text{SI}(SS \ l, AS \ i, 0).
\]

Again, \( \lambda_y \) is a normalization factor to make sure that \( \text{SI}(AS \ i, SS \ j, 0) + \text{SI}(AS \ i, SS \ j, 1) = 1 \).

- **Convergence check:** At the end of each iteration, we calculate the likelihood of each codeword bit. The codeword bit between \( AS \) \( i \) and \( SS \) \( j \) is temporally determined to be 1 if

\[
\text{SI}(AS \ i, SS \ j, 1) - \text{SI}(SS \ j, AS \ i, 1) > 0,
\]

and determined to be 0 otherwise. If the entire temporal decisions meet all the local constraint rules, then the algorithm is considered to be convergent. Otherwise, the algorithm restarts with another random initial priori probability.

IV. NUMERICAL RESULTS

The performance of a scheduling algorithm can be evaluated in terms of the throughput and the average packet delay. The throughput is defined as the average number of successfully transmitted packets per time slot, expressed by (3). For the average packet delay, we adopt a model developed by [4], which assumes independent Poisson transmissions of the packets. In that case, the average packet delay can be calculated using the Pollaczek-Khintchin equation [4]. For comparisons, we also consider other heuristic scheduling algorithms, including 1) the spatially opportunistic algorithm (SOA) [5], 2) the round-robin algorithm (RRA), and 3) the optimal exhaustive search algorithm. The optimal exhaustive search algorithm tries to maximize the total channel utilization by examining all the possible combinations. That is,

\[
\max \sum_{s=1}^{5} \sum_{k \in \{0,1\}} f'_s,
\]
where the set \( \{\chi\} \) is the orthogonal set, which is defined as follows. For any \( i \) and \( j \) belonging to the set \( \{\chi\} \), and \( i \neq j, i \) and \( j \) have to satisfy
\[
\chi_i \cdot \chi_j = 0 ,
\]
where “ \( \cdot \) ” is the inner product operator. This means that the users in the set \( \{\chi\} \) can be transmitted simultaneously without severe multiuser interference at the \( n \)th frequency sub-carrier. It should be noted that if there is no orthogonal set available, the optimal exhaustive search algorithm directly employs the SOA to schedule the resources. It should be noted that since the total channel utilization of the optimal exhaustive search algorithm is the optimal, therefore, the performance of the optimal exhaustive search algorithm can be regarded as the overall performance bound. Regarding the simulation setup, the first step in setting up the simulation is to generate each subscriber’s channel response in binary-valued 2-D CSI LUT. However, each subscriber’s binary-valued 2-D CSI LUT will be mainly dominated by two parameters. The first is the number of subscribers in the system. A system with a large number of subscribers will experience more interference than one with a smaller number of subscribers. The second is the pre-determined threshold for establishing a reliable link. This will depend on the required SINR for an individual subscriber’s service demand. In our simulation scenario, four subscribers need to be scheduled within the AP service area, where the AP has five beams and eight frequency sub-carriers as the total system resources. In this case, we randomly generate each subscriber’s channel response structure with binary-valued 2-D CSI LUT based on the channel model suggested in [2,6], where the mean and variance of each subscriber’s channel utilization is set up to 33.63% and 1.05%, respectively, which are based on above simulation experiment scenario. Also, we assume that, within a time slot, each subscriber’s channel statistical behavior will remain the same. That is, the steering vectors of the angles-of-departure (AODs) at AP and that of the angles-of-arrival (AOAs) at the subscribers usually remain constant over hundreds of data bursts, according to the channel structure characteristics in [6].

The average total channel utilization evaluated over 10,000 time slots for the RRA, the SOA, the SOA, the proposed normal graph algorithm (NGA), and the optimal exhaustive search algorithm are 33.42%, 53.97%, 63.54%, and 71.62%, respectively. We can see that the proposed normal graph-based algorithm performs worse than the optimal exhaustive search algorithm but better than all the other algorithms. In the following, the performance metric with which we are concerned is the average packet delay. The average packet delay versus the total arrival rate for four scheduling algorithms is shown in Fig. 3 in the revised manuscript. As shown in Fig. 3, the total system delay experienced in the schedule from using the proposed algorithm is worse than that from using the optimal exhaustive search algorithm and much less than those from using both the RRA algorithm and that by the SOA algorithm. Finally, let us consider the issue of algorithm complexity. Since the operations of the compared algorithms are completely different, it is difficult to compare the algorithm complexity in term of computation complexity. However, if we compare the algorithm complexity from the viewpoint of individual algorithm complexity versus the involved users \( K \) for each sub-carrier \( n \), we can see that 1) the RRA algorithm is linearly proportional to \( K \), 2) the SOA algorithm is proportional to \( K(\bar{K}−1)/2 \), 3) the proposed algorithm is also linearly proportional to \( K \), and 4) the optimal exhaustive search algorithm is proportional to \( \sum_{i=2}^{K} C_i^K \). It is obvious that the optimal exhaustive search algorithm is not suitable for implementation as the number of involved users is large. On the other hand, the individual algorithm complexity of the RRA and the proposed algorithm is less sensitive to the number of involved users, but the channel utilization and the average packet delay resulting from the use of the proposed algorithm are better than those of the RRA.

![Fig.3 Average time delay for different schemes.](image)

V. CONCLUSIONS

This paper proposes a low-complexity algorithmic framework for solving the MU-MIMO scheduling problem based on the modeling and computational methodology of normal graphs. In the proposed algorithm, soft information is exchanged among angle sub-channel nodes and subscriber nodes in a normal graph to iteratively optimize the spatial resource allocation. The performance of the proposed scheduler is compared to other two commonly-used schedulers, the SOA and the RRA, and the optimal exhaustive search algorithm in terms of the channel utilization and the average packet delay. The effectiveness of the proposed scheduler is verified through computer simulations under various channel scenarios.

REFERENCES


