Downlink Scheduling with Optimal Antenna Assignment for MIMO Cellular Systems

Young-June Choi, Jongtack Kim, and Saewoong Bahk
School of Electrical Engineering & Computer Science, Seoul National University
{yjchoi, jkim, sbahk}@netlab.snu.ac.kr

The bandwidth-limited wireless channel can considerably improve the performance by exploiting multiple-input–multiple-output (MIMO) antennas. Combining spatial multiplexing with multiuser diversity, we develop an optimal cross-layer scheduling mechanism that executes fair scheduling at the upper layer and optimal antenna assignment at the physical layer. For fair scheduling, we propose a framework that achieves the objective of maximum capacity and proportional fairness. For optimal antenna assignment, we consider the Hungarian algorithm that maximally utilizes the characteristics of MIMO systems by adopting the graph theoretical approach. Through simulations, we demonstrate the performance of optimal scheduling.

1. Introduction

A multiple-input–multiple-output system can increase the data rate by sending independent information streams over several transmit–receive antenna pairs simultaneously. A theoretical approach has shown that the capacity of a MIMO system linearly increases with the minimum number of transmit or receive antennas, which is represented as \(\min(m, n)\) and called the multiplexing gain \([2]\). Some previous works have designed efficient methods to select antennas by spatial multiplexing considering the minimum error rate \([3]\) and multiuser diversity \([4,6]\) depending on the physical layer technology.

The multiuser diversity is an efficient way to maximally utilize spatial multiplexing because each path experiences Rayleigh fading. As packet scheduling selects antennas according to the multiuser diversity, we combine antenna assignment with downlink packet scheduling from the viewpoint of cross-layer design between physical and network layers. The downlink requires efficient scheduling because the traffic for the downlink is much higher than that for the uplink. Hence we suggest an optimal antenna assignment mechanism to ensure fairness at the downlink while exploiting multiuser diversity.

The approach in \([5]\) uses round robin scheduling after selecting a set of users in advance which number is the same as that of transmit antennas. As it determines the user set without considering the channel state, it cannot fully use multiuser diversity. \([8]\) uses the proportionally fair scheduler for the \(2 \times 2\) MIMO system and compares the performance of centralized scheduler with that of two distributed schedulers. Our proposed scheduler runs in a centralized way that performs whole or partial allocation for the given set of transmit antennas. The centralized decision is beneficial to optimal scheduling in the aspect of efficient resource management.

In this paper we consider a downlink scheduling scheme for a MIMO cellular system by taking the approach of cross layer design and introduce a graph theoretical approach to solve the optimization problem. We perform optimal antenna assignment for spatial multiplexing by Hungarian algorithm. Then we apply the optimal assignment for fair scheduling of best-effort traffic.

We organize the remainder of this paper as follows. Section 2 illustrates the system model and Section 3 formulates three objectives for optimal antenna assignment and fairness. Section 4 develops an optimal solution for the scheduling. Section 5 shows the simulation results for the optimal assignment algorithms, followed by conclusions in Section 6.

2. System Model

We consider the downlink of a MIMO cellular system with \(m\) transmit and \(n(\geq m)\) receive antennas. We assume the base station (BS) communicates with \(k\) mobile users.

In current packet cellular networks such as HSDPA (3GPP) and cdma2000 1x EV-DO (3GPP2), the scheduler at the BS uses the
feedback channel information to know the exact channel condition for each user. Defining $h_{ij}$ as the fading coefficient between receive antenna $i$ and transmit antenna $j$, we can write a matrix $H = [h_{ij}] \in C^{m \times n}$. Supposing $h_{ij}$ is quasi-stationary and time-invariant during transmission of a packet, we can write the received signal as

$$Y = \sqrt{\frac{SNR}{m}} HX + N$$

(1)

where matrices $H$ and $X$ represent transmitted and received signals for multiple antennas respectively, and $N$ is a Gaussian random noise vector. SNR is the average signal to noise ratio at receive antennas.

To estimate the original signal $X$, a receiver should use $m \times n$ matrix equalizer $G$ such that

$$\hat{X} = GY$$

The channel capacity (bps/Hz) of MIMO systems with $m$ transmit antennas under AWGN is known as [1]

$$C(SNR) = \log_2 \det (I + \frac{SNR}{m}HH^*)$$

(2)

where $H^*$ represents the Hermitian matrix of $H$. Modifying (2), we can rewrite it for the spatial multiplexing with linear receiver as [3]

$$C = \sum_{i=1}^{m} \log_2 (1 + SNR_i)$$

(3)

where

$$SNR_i = \frac{E[|GH|^2]}{mN_0 \sum_j |G|^2 + E_s \sum_{j=1, j \neq i}^m |GH|^2}$$

(4)

Here $E_s$ is the received power and $N_0$ is the noise power per receive antenna.

For an MMSE (minimum mean-square error) receiver, $G$ is given by [3]

$$G = \left[H^*H + \frac{N_0}{E_s} I_m \right]^{-1} H^*$$

(5)

where $I_m$ is $m \times m$ identity matrix. In this paper, we assume that the channel is perfectly known at the receiver and the number of receive antennas is the same as that of transmit antennas, that is, $m = n$. Fig. 1 shows our MIMO channel model with $m$ transmit and receive antennas which consists of $m$ channels in conjunction with each transmit antenna.

3. Problem Formulation

In spatial multiplexing, a data stream is split into multiple substreams, and each substream is transmitted through a transmit antenna. To use multiuser diversity effect, the scheduler can simultaneously choose multiple users as many as transmit antennas and allocate a transmit antenna to a user. In this case it needs to find appropriate objective value (weight) per antenna for a user’s transmission, which reflects the channel condition estimated by the measured SNR. From now on, we assume that each user always has packets (i.e., active) to receive from the BS.

3.1 Scheduler-I

As the capacity is a function of SNR, maximizing the sum of weights obtained from SNR is equivalent to the capacity (or throughput) maximization of (3). That is, find $k_1, \ldots, k_m \in K$ for $\forall i, j (\neq i) k_i \neq k_j$

$$\text{s.t. } \max C(k_1, \ldots, k_m) = \max \sum_{i=1}^{m} \log_2 (1 + SNR_i(k_i))$$

(6)

where $C(k_1, \ldots, k_m)$ is the capacity achieved by allocating transmit antenna $i$ to user $k_i$ for $i \in \{1, \ldots, m\}$. $K$ is a set of candidate users controlled by fairness. This scheduler provides unfair service because its primary goal is to maximize capacity. We will consider the fairness issues in Section 3.3.

We first choose $m$ users for $m$ transmit antennas assuming the number of users $k \geq m$, and consider the case of $k < m$ at the
end of next section. The complexity of exhaustive search for the optimal allocation is directly related with \( \binom{k}{m} \) which exponentially increases with the increase of \( k \). Therefore we present a solution to find the optimal multiplexing assignment with low complexity. Solving (6) with low complexity, we are also able to take the same approach for the next scheduling strategy.

### 3.2 Scheduler-II

Proportionally fair scheduling has been known as an appropriate criterion for best-effort traffic [8,9] though it does not maximize throughput. Extending the proportionally fair scheduling in [7] to MIMO systems, we can choose the best \( m \) users as follows.

\[
\text{Find } k_1, \ldots, k_m \in K \text{ for } \forall i, j (\neq i) \quad k_i \neq k_j \\
s.t \quad \max U(k_1, \ldots, k_m) = \max \sum_{i=1}^{m} r_i(k_i) 
\]

where \( r_i(k_i) \) and \( R_i(k_i) \) are the instant and the average data rate of user \( k_i \) respectively measured in a channel between transmit antenna \( i \) and user \( k_i \). At time slot \( t+1 \), the average data rate of user \( k_i \) for transmit antenna \( i \) is updated by [7]

\[
R_i(k_i, t+1) = \begin{cases} 
(1-1/t_c) \cdot R_i(k_i, t) + 1/t_c \cdot r_i(k_i, t) & \text{if } k_i \text{ is selected for transmit antenna } i, \\
(1-1/t_c) \cdot R_i(k_i, t) & \text{otherwise}
\end{cases}
\]

where \( t_c \) is the averaging factor.

### 3.3 Fairness

We can achieve fairness by two kinds of approaches, weight-based fairness and frame-based fairness. According to the design goal, fairness can have various objectives such as throughput-fair and slot-fair. While a slot denotes a unit resource used by one antenna, a frame indicates a large-scale unit consisting of multiple slots. If the length of a frame is \( T \) (slots), the total number of slots in a frame is \( mT \) since there are \( m \) transmit antennas. For simplicity, we consider only the slot-fair scheduling for the two approaches.

The frame-based fairness is to ensure fair scheduling at the frame level. If a scheduler adopts slot-fair scheduling, the slots in a frame should be distributed evenly to all users. Letting \( \phi_i \) be the user \( i \)'s weight for fairness, and \( L \) be the reserved number of slots for real-time services. The target number of slots for user \( i \) based on the slot-fair scheduling is given by

\[
\left\lfloor \frac{mT-L}{\sum_{j=1}^{K} \phi_j} \right\rfloor \tag{9}
\]

Within a frame, users join the candidate set \( K \) for current scheduling unless the target number is 0, while the target number decreases by 1 if the user is scheduled. The frame-based fairness is suitable for the scheduler-I in (6).

Meanwhile, in the weight-based fairness, the weight value itself in the multiplexing reflects fairness. Proportional fairness mentioned in the scheduler-II is a good example of weight-based schemes. As time goes to infinity, the proportionally fair scheduling approximately has the same result as the slot-fair scheduling with \( \phi_i = \phi \) when channel fading is identical and independent. This means that users will have the same number of slots on the average in the long run. Another example of weight-based fairness is introduced in [9], which develops a general scheduling policy in channel-varying wireless links. Though it can not maximize the instant throughput, it has an advantage of supporting fairness without using any frame structure.

### 4. Optimal Solution

Now we solve the optimization problem that finds the user-antenna matching for the above objectives, taking a graph theoretical approach.

#### 4.1 Graph theoretical solution

For a graph theoretical solution, we assume every weight has a nonnegative real number. We take two vertex sets – one for users and the other for transmit antennas. Fig. 2 shows two sides of vertex sets for the case of 4 users and 4 transmit antennas. We draw an edge to represent a channel between a user
vertex and an antenna vertex. The edge is with some weight to indicate the SNR in (6) or the fairness index in (7). Thus the original optimization problem is equivalent to the optimal assignment problem applicable to weighted bipartite matching shown in Fig. 2.

The optimal assignment problem is to assign \( n \) jobs to \( n \) workers. Each worker who is qualified for one or more jobs and has different efficiency for each job is assigned to a job to maximize the total efficiency. This problem was solved by Hungarian algorithm using bipartite matching graphs [10,11].

### 4.2 Hungarian algorithm

Now we describe the procedures of Hungarian algorithm. Matrix \( \mathbf{W} = [w_{ij}] \) has the element of weight \( w_{ij} \) for antenna \( i \) and user \( j \) as shown in Fig. 3 (a).

1) Step 1: Let \( X, Y \) be the bipartition sets. Initialize two labels \( u_i \) and \( v_j \) by

\[
u_i = \max_{j} w_{ij}, \quad v_j = 0, \quad i,j = 1, \ldots, k.
\]

In Fig. 3 (b), the numbers written at the left and top of each matrix express \( u_i \)'s and \( v_j \)'s respectively.

2) Step 2: Obtain the excess matrix \( \mathbf{C} \) by the following.

\[
c_{ij} = u_i + v_j - w_{ij}
\]

3) Step 3: Find the subgraph \( G \) that includes vertices \( i \) and \( j \) satisfying \( c_{ij} = 0 \) and the corresponding edge \( c_{ij} \). Then find the maximum matching \( M \) in \( G \) and draw underline for the entries in \( M \). If \( M \) is perfect matching with \( k \) edges, go to step 5.

4) Step 4: Let \( Q \) be a vertex cover of \( G \), and let \( R = X \cap Q \) and \( T = Y \cap Q \). The vertex cover \( Q \) as a vertex set of \( G \) contains at least one endpoint of each edge. Now find \( \epsilon \) satisfying

\[
\epsilon = \min \left\{ w_{ij} : x_i \in X - R, \ y_j \in Y - T \right\}.
\]

For instance \( \epsilon \) equals 1 in Fig. 3. Decrease the \( u_i \) by \( \epsilon \) for the rows of \( R^c \) and increase \( v_j \) by \( \epsilon \) for the columns of \( T \). Then go to step 2.

5) Step 5: \( M \) is the optimal assignment solution when \( M \) is perfectly matched with \( k \) edges [12].

### 4.3 Remark

Theorem 1: The Hungarian algorithm terminates within \( n^2 \) iterations in the \( n \times n \) bipartite graph.

Theorem 2: The Hungarian algorithm finds a maximum weight matching.

The proofs of Theorem 1 and 2 are given in [12]. By using these, we can apply Hungarian algorithm for the optimal antenna assignment. However, the solution is optimal only when the cardinalities of both vertex sets in a bipartite matching graph are the same. To relax the condition of relation between \( k \) and \( m \), we need the following theorem.

Theorem 3: \( k \times m (k > m) \) assignment is equivalent to a \( k \times k \) problem when \( k - m \) null vertices are inserted.

In case of \( k < m \), we have more transmit antennas than the number of users \( k \). So some users can be allocated for more than one transmit antennas. Hungarian algorithm is also applicable by allowing these users to have two or more vertices.

### 5. Simulation Results

In the 4 \( \times \) 4 MIMO system, we evaluate the performance of optimal scheduling for best-effort packets. We consider a downlink channel and the cell radius is 1 Km. The channel is influenced by path loss with exponent 4, slow fading with log standard deviation 8 dB, and Rayleigh fading with zero
mean and unit variance. One slot time is 2 msec and the averaging factor \((t_c)\) in (8) is 1000 slots.

Fig. 4 depicts the capacity comparison among round robin (RR), antenna-assisted round-robin (AARR) in [5], frame-based slot-fair scheduling (FF) of (6), proportionally fair scheduling (PF) of (7) where the last two are performed by Hungarian algorithm for optimality. It shows that PF with Hungarian algorithm also utilizes the capacity maximally. Even the capacity over 10 users is better than that of FF. This is because the FF is limited by the frame structure. On the other hand, the performance of AARR is relatively low because it performs round robin and tries optimal scheduling only for the pre-determined \(m\) users.

6. Concluding Remarks

In this paper, we decoupled the optimal scheduling problem under MIMO environments into the fair scheduling at the upper layer and the optimal antenna assignment at the physical layer. Combining spatial multiplexing with multiuser diversity, the scheduler chooses multiple users as many as the number of transmit antennas in order to maximize capacity or to support proportional fairness. We solve the general optimal multiplexing problem by Hungarian algorithm that requires low complexity. Simulation results show that the optimal assignments increase the capacity and proportional fair scheduling can also be a good choice for MIMO systems. As a user requires different service quality according to its service class, we developed a framework of optimal packet scheduling to accommodate heterogeneous services by jointly using MIMO diversity and spatial multiplexing.

Reference


