Research Article

Efficient Wireless Broadcasting through Joint Network Coding and Beamforming

Monchai Lertsutthiwong, Thinh Nguyen, and Bechir Hamdaoui

School of EECS, Oregon State University, Corvallis, OR 97331, USA

Correspondence should be addressed to Thinh Nguyen, thinhq@eecs.oregonstate.edu

Received 19 July 2011; Accepted 18 November 2011

Academic Editor: Manzur Murshed

Copyright © 2012 Monchai Lertsutthiwong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We develop a framework that exploits network coding (NC) and multiple-input/multiple-output (MIMO) techniques, jointly together, to improve throughput of downlink broadcast channels. Specifically, we consider a base station (BS) equipped with multiple transmit antennas that serves multiple mobile stations (MSs) simultaneously by generating multiple signal beams. Given the large number of MSs and the small number of transmit antennas, the BS must decide, at any transmission opportunity, which group of MSs it should transmit packets to, in order to maximize the overall throughput. We propose two algorithms for grouping MSs that take advantage of NC and the orthogonality of user channels to improve the overall throughput. Our results indicate that the proposed techniques increase the achievable throughput significantly, especially in highly lossy environments.

1. Introduction

In recent years, multiple-input/multiple-output (MIMO) has been recognized as a key enabling technology for improving the performance of wireless communication systems. Unlike traditional communications, MIMO techniques rely on multiple antennas to transmit and/or receive signals. The number of antennas that a device can be equipped with can be limited due to spatial constraints. For example, in a cellular network, while there may be no limit on the number of antennas that the base station (BS) can be equipped with, there is a limit on that number when it comes to mobile station (MS) due to size and/or cost constraints. MIMO capabilities can be exploited to enable the spatial division multiple access (SDMA) technique, which basically allows multiple simultaneous transmissions from the BS to multiple MSs, thereby achieving higher overall data throughput [1–3]. Specifically, the BS exploits MIMO to generate radiation patterns that simultaneously target different groups of MSs [4, 5]; this is known as beamforming (BF). Mathematically, the signal beam from each antenna is coded and multiplied independently by a BF weight vector to control its shape and direction. A BF weight vector is determined by MSs locations as well as the characteristics of channels between the BS and MSs. Assuming that the BF weight vectors are available, an optimal SDMA scheme consists then of selecting the set of MSs whose signal-to-interference plus noise ratios (SINRs) are maximized.

A BS with $M$ transmit antennas can form and transmit a signal with at most $M$ beams [6, 7]. Therefore, assuming that one beam is allocated for each MS, when the number of MSs, $K$, exceeds the number of transmit antennas, $M$, the BS selects the best $M$ MSs, that is, those whose SINRs are maximized [8]. An efficient scheduling algorithm is then required at the BS to map the MSs with these beams. For example, given 10 MSs ($z_1, z_2, \ldots, z_{10}$) and one BS equipped with 2 antennas, an efficient scheduler may select ($z_2, z_4$) in the first time slot for transmission, ($z_1, z_5$) in the second time slot, ($z_7, z_9$) in the third time slot, and so on. After five time slots, the BS completes then transmitting all packets to their intended MSs. The key challenge in designing these scheduling algorithms lies in grouping the MSs in such a way that interference among MSs is minimized; that is, a transmitted beam intended for one MS should create no or minimal interference to any other MSs in the group. One way to achieve this objective is to group together MSs whose channel vectors are somewhat orthogonal to one another.
Network coding (NC) is another effective technique that is also well recognized for its great network throughput potentials and has received a considerable attention due to its practical simplicity [9]. In single-hop wireless networks, such as cellular networks, NC can be used to significantly reduce the number of transmission/retransmission attempts required to successfully deliver data packets [10]. Consider the example of Figure 1 for the purpose of illustration.

Instead of immediately retransmitting a lost packet for the purpose of illustration, NC can be used to significantly reduce the number of transmission/retransmission attempts [10]. Consider the example of Figure 1 for the purpose of illustration.

![Figure 1: Illustrative example of NC technique.](image)

**Figure 1:** Illustrative example of NC technique.

Network coding (NC) is another effective technique that is also well recognized for its great network throughput potentials and has received a considerable attention due to its practical simplicity [9]. In single-hop wireless networks, such as cellular networks, NC can be used to significantly reduce the number of transmission/retransmission attempts required to successfully deliver data packets [10]. Consider the example of Figure 1 for the purpose of illustration.

Instead of immediately retransmitting a lost packet for the purpose of illustration, NC can be used to significantly reduce the number of transmission/retransmission attempts [10]. Consider the example of Figure 1 for the purpose of illustration.

![Figure 1: Illustrative example of NC technique.](image)

**Figure 1:** Illustrative example of NC technique.

Joint BF-NC-based scheduling algorithms. We show the simulation results for the proposed techniques in Section 6 and conclude the paper in Section 7.

### 2. Related Work

Simultaneous service support of multiple MSs through BF techniques can be very challenging due to interference incurred by different active signal beams. Therefore, several efficient scheduling algorithms and proper BF strategies have been proposed to reduce interference among multiple beams. To maximize the sum capacity, the scheduled MSs are typically selected through an exhaustive search, which can be very costly or even infeasible for a large MS pool. For example, Shad et al. [11] proposed an algorithm for selecting the MSs for each time slot over SDMA/TDMA networks by considering the largest SINR margin of scheduled MSs with respect to some SINR thresholds. Choi et al. [12] provided a scheduling algorithm, which relies on the feedback from MSs during each time slot to determine a subset of MSs in such a way to maximize the sum capacity. Ravindran et al. [13] studied the performance of zero-forcing BF techniques under limited channel feedback in line of sight (LOS) channels. Their results show that a zero-forcing BF technique performs well especially when the system has reliable channel state information (CSI) from MSs. Furthermore, Yoo and Goldsmith [7] considered the orthogonality of CSIs among the MSs to schedule the MSs while minimizing the overall interference. In particular, their technique greedily selects a set of MSs while assuming that the BS knows CSIs of all MSs in the system.

Another traditional technique, yet efficient, is to use a BF codebook containing a set of all possible coefficients (or codewords) to form a BF beam. In general, the codebook is generated offline and consists of a fixed number of codewords, which do not depend on statistical characteristics of channel conditions [14]. Most codebooks are fixed and designed for a specific transmitter. For example, Love and Heath [15] proposed a way to generate a codebook that cannot adaptively change as the channel changes. By using their technique, both BS and MSs have full knowledge of a codebook. Amiri et al. [16] provided an efficient algorithm to adaptively change a codebook based on new feedback information at the BS. However, their technique is complex because the BS needs to frequently evaluate new feedback information, and hence, the technique may not be suitable for systems with a large number of MSs. Huang et al. [17] showed the performance of orthogonal BF techniques by selecting a codeword for each MS in such a way to reduce the interference among all active beams. More specifically, their technique consists of finding a set of codewords from a fixed codebook for all active beams by relying on MSs CSIs.

Most of these reported works consider associating/allocating one BF beam for each intended MS. However, MSs within a close proximity often experience similar channel conditions and have similar SINR levels. Therefore, considering grouping and associating these MSs together under one BF beam not only allows to support more MSs but also enables effective use of NC techniques, thus improving
throughput even further. This work proposes a framework that takes into consideration the proximity in SINRs and CSI orthogonality when grouping MSs and forming BF beams. We first prove that the group formation problem is NP-hard (the proof is given in the appendix) and then propose two heuristic techniques: a greedy-selection-based technique and a codebook-based technique, which are described later in Section 5.

3. Network Coding and Broadcasting

For notational convenience, throughout this paper, uppercase boldface letters are used for matrices and lowercase boldface letters are used for vectors, and \( X^\ast, X^T, \tilde{X}, |X|, \) and \( ||X|| \) denote conjugate transpose, transpose, normalized, size or magnitude, and L-2 norm of \( X \), respectively.

In this section, we briefly describe the use of NC technique over wireless broadcast channels. We consider \( K \) MSs (as receivers) and one BS (as the sender). Without loss of generality, we assume that each data packet can be transmitted in one time slot (all time slots are of a fixed duration). We also assume that MSs use either positive or negative acknowledgements (ACK/NACK) to let the BS know about the status of receiving each packet. Traditionally, an MS immediately sends a NAK if the packet is not successfully received. When using an XOR NC technique, instead of immediately retransmitting a lost packet to a certain MS right after receiving a NAK, the BS waits for a fixed interval of time (in time slots) before retransmitting starts. With NC enabled, the BS then retransmits an XOR-coded packet for each set of multiple lost packets. For this to happen, the BS maintains a list of all lost packets as well as their corresponding MSs for some time (i.e., waiting period) during retransmission phase. The number of lost packets that can be coded at once is then at most equal to the number of MSs, \( K \), one lost packet from each MS. Any successful packets received by any of the \( K \) MSs will not be included in the coded packet to be retransmitted. The BS keeps retransmitting coded packets of lost packets until all MSs receive all packets successfully.

Consider a broadcast transmission example of a cellular network, where all BS packet transmissions are assumed to reach all MSs. For a system with two MSs, that is, \( z_1 \) and \( z_2 \), we consider the pattern of lost packets shown in Figure 2 as an example, where a cross \( X \) and an oval \( O \), respectively, represent a lost and a coded packet. Since \( K = 2 \), the coded packets generated by the BS and intended for the two MSs, \( z_1 \) and \( z_2 \), are \((a_1 + a_2), (a_3 + a_4), \) and \( a_6 \). Packet \( a_6 \) is only intended for MS \( z_2 \) because MS \( z_1 \) has already received it successfully. Note that if some coded packets cannot be correctly received by any of the intended MSs, the BS will update the lost packet table and determine the best coded packet for the next retransmission opportunity.

Without loss of generality, for a system with \( K \) MSs, we assume that \( \text{PER}_i \geq \text{PER}_j \) if \( i \leq j \) for any \( i, j \in \{1, 2, \ldots, K\} \) where \( \text{PER}_i \) is the packet error rate for MS \( z_i \). Recall that the BS will generate a coded packet combining as many lost packets as possible, and the MS with the highest \( \text{PER} \) dominates the performance of the system. Let \( z_K \) be this MS.

Let \( X_i \) be a random variable representing the number of lost packets of MS \( z_i \) after a certain number \( N \) transmissions. When \( N \) approaches infinity and each transmission follows a Bernoulli trial, \( X_i \) will follow the Binomial random variable. Therefore, we have

\[
E \left[ \max_{i \in \{1, 2, \ldots, K\}} X_i \right] \approx E[X_K].
\]

With \( N \) transmitted packets to \( K \) MSs, the expected number of transmissions can be written as

\[
N + E[X_K] = N + \frac{N \times \max_{i \in \{1, 2, \ldots, K\}} \{\text{PER}_i\}}{1 - \max_{i \in \{1, 2, \ldots, K\}} \{\text{PER}_i\}}.
\]

Note that the NC technique described in this section considers that packets will be broadcast to all MSs. Thus, packets sent by the BS can be heard either by all MSs or by none.

4. Joint Beamforming and Network Coding

In this work, we consider broadcast transmissions in a downlink channel of a cellular network, where data packets sent by the BS are intended for and reach all MSs in the network. When data packet transmission is not successful, the BS keeps retransmitting until the lost/unsuccessful packet is delivered to all MSs successfully. We assume that the BS needs to broadcast \( D \) packets to the MSs.

After the first transmission, one or more MSs may not receive the packets. When this happens, one simple way to ensure successful delivery of packets is for the BS to retransmit/broadcast each lost packet to all MSs again (e.g., Figure 3(a)). Another retransmission technique is to use beamforming to retransmit the lost packets only to those MSs that did not receive the packets. In this paper, we investigate and develop techniques that rely on MIMO to form BF beams so as to direct transmission towards a set of MSs only, that is, those MSs that did not receive the packets (e.g., Figure 3(b)). In addition to BF techniques, our proposed techniques rely on and incorporate network coding (NC) to further improve throughput performances of broadcast channels.
Simply broadcasting lost packets to all MSs, instead of using BF technique, can be very inefficient for the following reasons. First, MSs are often located at different positions with various distances from the BS, leading to high variations of MSs received signal strengths. With this, different MSs are likely to experience different packet success rates. Second, without directing the transmitted signals, power could be wasted when signals are sent in directions/regions that happen to have no MSs. Let us consider the example shown in Figure 3, where there are 3 groups of MSs: \{z_1, z_2\}, \{z_3\}, and \{z_4\}. In this example, transmitting packets to a location that does not contain any of the groups will be a waste of energy resources. By using a BF technique as in Figure 3(b), the BS can direct BF beams in such a way to increase the signal quality reaching the intended MSs. Note that the BS may use multiple time slots for transmitting a packet to everyone if all MSs are located apart from each other. The challenge as well as the novelty of the proposed BF technique lies in grouping all MSs into multiple clusters each corresponding to one beam in such a way to enhance broadcast throughput while minimizing power consumption.

We now describe how to combine NC and BF to further improve broadcast network throughput. Figures 4(a) and 4(b) illustrate the use of NC to broadcast a stream of packets in the network, respectively, shown in Figure 3(a) (without beamforming) and Figure 3(b) (with beamforming). In this example, there are 4 MSs (i.e., \{z_1, z_2, z_3, z_4\}). The BS needs to deliver 7 packets (i.e., \{a_1, a_2, ..., a_7\}) to all MSs and can generate at most 3 beams at once. Recall that the best strategy to generate a coded packet is to maximize the number of lost packets that can be coded at once. Therefore, we use 3 time slots to generate and send coded packets. In the first time slot, when not using a BF technique, the coded packet to be sent to all MSs is \((a_1 \oplus a_2 \oplus a_3 \oplus a_4)\), whereas, when using a BF technique, the three coded/uncoded packets, \((a_1 \oplus a_2)\), \(a_3\), and \(a_4\), are to be sent on beams 1, 2, and 3, respectively. In the second time slot, the BS generates coded packets that combine the lost packets \(a_5\) and \(a_6\) and sends them to MSs \(z_1\), \(z_2\), and \(z_4\). When not using a BF technique, the coded packet \((a_5 \oplus a_6)\) is then to be sent to all MSs. Whereas, when using a BF technique, the coded packet \((a_5 \oplus a_6)\) and uncoded packet \(a_6\) are to be transmitted through beams 1 (for \(z_1\) and \(z_2\)) and 3 (for \(z_4\)), respectively. In the third time slot, the BS needs to transmit only packet \(a_7\) to MS \(z_4\). In this case, it is obvious that there is no need to broadcast a packet \(a_7\) to everyone. Hence, when using a BF technique, the signal quality sent to MS \(z_4\) can be improved by allocating all transmit power to beam 3, thereby increasing the packet success rate and/or reducing energy consumption.

In the first time slot, note that the coded packet combining the first 4 lost packets (i.e., without BF) when sent to all MSs exactly suffices as it enables each of the 4 MSs to recover its lost packet (i.e., \(a_i\) for \(z_i\)). Thus, this may seem to be the best strategy. Likewise, the BF technique also suffices; that is, sending \((a_1 \oplus a_2), a_3\), and \(a_4\) on three different beams 1, 2, and 3, respectively, also enables all MSs to recover their lost packets. The difference between the two techniques lies, however, on the fact that the BF technique can make received signals stronger, thus decreasing the packet error rates by increasing the chance of correctly decoding received packets.

In Figure 3(b), the BS can direct BF beams in such a way to enhance broadcast throughput while minimizing power consumption. Figures 4(a) and 4(b) illustrate the use of NC to broadcast a stream of packets in the network, respectively, shown in Figure 3(a) (without beamforming) and Figure 3(b) (with beamforming). In this example, there are 4 MSs (i.e., \{z_1, z_2, z_3, z_4\}). The BS needs to deliver 7 packets (i.e., \{a_1, a_2, ..., a_7\}) to all MSs and can generate at most 3 beams at once. Recall that the best strategy to generate a coded packet is to maximize the number of lost packets that can be coded at once. Therefore, we use 3 time slots to generate and send coded packets. In the first time slot, when not using a BF technique, the coded packet to be sent to all MSs is \((a_1 \oplus a_2 \oplus a_3 \oplus a_4)\), whereas, when using a BF technique, the three coded/uncoded packets, \((a_1 \oplus a_2)\), \(a_3\), and \(a_4\), are to be sent on beams 1, 2, and 3, respectively. In the second time slot, the BS generates coded packets that combine the lost packets \(a_5\) and \(a_6\) and sends them to MSs \(z_1\), \(z_2\), and \(z_4\). When not using a BF technique, the coded packet \((a_5 \oplus a_6)\) is then to be sent to all MSs. Whereas, when using a BF technique, the coded packet \((a_5 \oplus a_6)\) and uncoded packet \(a_6\) are to be transmitted through beams 1 (for \(z_1\) and \(z_2\)) and 3 (for \(z_4\)), respectively. In the third time slot, the BS needs to transmit only packet \(a_7\) to MS \(z_4\). In this case, it is obvious that there is no need to broadcast a packet \(a_7\) to everyone. Hence, when using a BF technique, the signal quality sent to MS \(z_4\) can be improved by allocating all transmit power to beam 3, thereby increasing the packet success rate and/or reducing energy consumption.

In the first time slot, note that the coded packet combining the first 4 lost packets (i.e., without BF) when sent to all MSs exactly suffices as it enables each of the 4 MSs to recover its lost packet (i.e., \(a_i\) for \(z_i\)). Thus, this may seem to be the best strategy. Likewise, the BF technique also suffices; that is, sending \((a_1 \oplus a_2), a_3\), and \(a_4\) on three different beams 1, 2, and 3, respectively, also enables all MSs to recover their lost packets. The difference between the two techniques lies, however, on the fact that the BF technique can make received signals stronger, thus decreasing the packet error rates by increasing the chance of correctly decoding received packets.

In Figure 3(b), the BS can direct BF beams in such a way to enhance broadcast throughput while minimizing power consumption.
Therefore, efficient grouping algorithms that can determine the best sizes/groups are needed. In this paper, we develop and propose two efficient algorithms for grouping MSs.

We assume that the BS is equipped with multiple antenna systems that enable it to simultaneously generate multiple transmitted beams. We also assume that the BS knows all channel vectors $\mathbf{h}_i$ associated with each MS $z_i$ for all $i$ through feedback channels. Let $P_i$ be the power allocated to beam $i$ by the BS. For each beam $i$, we assign a group of MSs $V_i$ where $V_i \subseteq V$, $V_i \cap V_j = \emptyset$ for any $i, j \in \{1, 2, \ldots, M\}$, and $V$ is a set of all MSs. Groups are formed in such a way that all MSs belonging to the same group are likely to experience the same packet losses, reducing then the number of transmissions at the BS via bitwise XOR operations.

The objective is to minimize the time $T$ for the BS to successfully deliver all $D$ packets to each MS. Note that minimizing the time $T$ is equivalent to maximizing the overall throughput under some fairness constraints. However, if a particular MS is far away from the BS such that it cannot receive packets, then clearly $T = \infty$. In this work, we do not consider this degenerated case, and such an MS is likely to be serviced by a neighboring BS. Furthermore, we assume that the BS uses a round-robin transmission style and the channels at MSs are stationary. Thus, each MS will receive a fair share of transmission packets and the expected achievable throughput would reflect the performance of the considered system.

Recall that the signal quality at MSs can be assessed/quantified via SINR. In particular, SINR$_{ij}$ for MS $z_j$ in beam $i$ can be computed as

$$\text{SINR}_{ij} = \frac{P_i |\mathbf{h}_i \mathbf{w}_i|^2}{\sum_{k=1}^{M} P_k |\mathbf{h}_k \mathbf{w}_k|^2 + \sigma^2},$$

where $P_i$, $\mathbf{w}_i$, $\mathbf{h}_i$, $\sigma^2$, and $M$ are transmit power, weight vector for beam $i$, channel vector for MS $z_j$, noise power, and number of beams, respectively. If all transmitted beams are pairwise orthogonal such that $|\mathbf{w}_i \mathbf{w}_k|^2 = 0$ and $|\tilde{\mathbf{h}}_i \mathbf{w}_i|^2 = 1$ (assume that $\mathbf{h}_i$ is normalized as $\tilde{\mathbf{h}}_i$ so as to check its orthogonality), where $z_j \in V_i \setminus V_k$, then $|\mathbf{h}_j \mathbf{w}_j|^2 = 0$. Thus, (3) would reduce to $\text{SINR}_{ij} = P_i |\mathbf{h}_i \mathbf{w}_i|^2 / \sigma^2$ (signal-to-noise...
ratio (SNR)). If all scheduled MSs have high SINR (or SNR), the equal power allocation is close to an optimal water-filling method [17]. Thus, the allocated transmit powers for all beams are all equal. Given a fixed total transmit power $P$ at the BS, this implies that the transmit power allocated for each beam $j$ is $P_j = P/M$ where $M$ is the number of active beams during that time.

Let $n_j$ be a set of lost packets for MS $z_j$. The number of transmissions for broadcast schemes over a group $V_i$ depends on the combination of all lost packets from all MSs in $V_i$. Therefore, a set of coded and uncoded lost packets based on the NC technique applied to beam $i$, $PER_{ij}$ is packet error rate for MS $z_j$ in beam $i$, and $x(i, j)$ is an assignment binary variable for MS $z_j$ in beam $i$ defined as

$$x(i, j) = \begin{cases} 
1 & \text{if MS } z_j \text{ is in beam } i, \\
0 & \text{otherwise}.
\end{cases}$$

Furthermore, each MS can only belong to one beam; that is, $\sum_i x(i, j) = 1$. For simplicity of analysis, we do not use any error correcting code. Therefore, $PER_{ij}$ can be computed in terms of bit error rate (BER$_{ij}$) such that $PER_{ij} = 1 - (1 - BER_{ij})^a$, where $a$ is a packet size in bits. Assume BPSK is used. Thus,

$$BER_{ij} = 0.5 \left[ 1 - \text{erf} \left( \frac{\sqrt{\text{SINR}_{ij}}}{\sqrt{10}} \right) \right],$$

where $\text{erf}(y) \approx (2/\sqrt{\pi}) \left[ y - (y^3/3) + (y^5/10) \right]$ and $y = \sqrt{\text{SINR}_{ij}}$ [18]. The expected number of transmissions for this system is the time that the BS spends until all MSs successfully receive all of their required packets.

Without loss of generality, (4) can be written as

$$\frac{N_{V_i}}{1 - \max_j \left[ x(i, j) \times PER_{ij} \right]},$$

where $l_q$ is the index of sets of BF beams. For each set $l_q$, the BS requires to successfully transmit all lost packets over a set of all BF beams before moving forward to the next set of BF beams, that is, $l_{q+1}$. Assume there are $Q$ sets of BF beams to be used at the BS. By taking into account all $Q$ sets of BF beams, our goal is then to minimize the total number of transmissions; that is,

$$\min \sum_q \left\{ \max_i \left\{ \frac{N_{V_i}}{1 - \max_j \left[ x(i, j) \times \text{PER}_{ij} \right]} \right\}, \right\},$$

s.t. $\sum_i x(i, j) = 1, \quad x(i, j) \in \{0, 1\},$

where $q = \{1, 2, \ldots, Q\}$. The solution to this optimization problem will be $x(i, j)$ assignments that minimize the total number of transmissions. However, we prove that this assignment problem is NP-hard (we provide the proof of NP-hardness in the appendix), and, instead, we propose two heuristics/techniques that solve this group formation problem by taking advantage of NC capabilities and the orthogonality of user channels. One proposed technique uses greedy selection based on SINR levels, whereas the other technique uses a predefined codebook to generate a set of orthogonal BF beams. Let $L = \{l_1, l_2, \ldots, l_Q\}$ denote the Q sets of $M$ orthogonal beams. For example, $l_i = 1$ means the BS uses the $i$th set of $M$ orthogonal beams in a codebook, otherwise $l_i = 0$. Next, we describe our two proposed scheduling techniques.

## 5. Proposed Scheduling Algorithms

Because our assignment problem is NP-hard, we propose to use two techniques to design heuristic algorithms that efficiently solve it. The first technique schedules and groups MSs based on their SINRs, and the second one uses a predefined codebook that first generates multiple sets of orthogonal BF beams and then assigns each MS to the beam whose SINR is maximized. The details of the proposed techniques will be described next.

### 5.1. Technique 1: Greedy-Selection-Based Scheduler

We first describe and formulate the group formation problem, which basically tries to find the optimal set of MSs that maximizes the overall throughput and then proposes a scheduling algorithm that solves it. The scheduling algorithm improves overall throughput through joint exploitation of both BF and NC techniques. Given an $M$-transmit antenna BS and $K$ one-receive antenna MSs located within the communication range of the BS, the BS can choose to transmit packets to a selected set of MSs that are located far apart from each other and whose quality levels are above a certain threshold. Note that the use of linear, equally spaced antenna arrays limits the number of concurrent signal beams to a certain number. We denote this number by $M_{\text{opt}}$, where $M_{\text{opt}} \leq M$.

When the number of MSs, $K$, exceeds $M_{\text{opt}}$, the BS chooses the best $M_{\text{opt}}$ (or fewer) MSs with the highest SINRs. We consider spatial separation among MSs by using the orthogonality in channel vectors, described in [7]. Below, we describe our greedy-selection-based scheduling algorithm.

The details of our proposed algorithm are described as follows: in step 1, $T_i$ is a set of all available MSs and
is a set of scheduled MSs where $i$ represents the current iteration number. Thus, $\delta$ is initially empty and $T_i$ is equal to $\{1, 2, \ldots, K\}$. In step 2, we compute $g_k$, the component of channel vector $h_k$ for each $k \in T_i$, orthogonal to previously determined subspace spanned by $\{g_1, \ldots, g_{i-1}\}$. Note that when $i = 1$, we assume $g_k = h_k$ for each MS $k$. In step 3, we select the MS, denoted by $\pi_i$, whose norm $\|g_i\|$ is maximized (i.e., best MS) over each $k \in T_i$. After that we update the values of $\delta$, $h_i$, and $g_i$, consecutively. In step 4, after finding the best MS $\pi_i$, we further determine a cluster of MSs whose SINRs are greater than some threshold $\Omega$. We denote this cluster by $\Omega_i$. Note that the value of $\Omega$ depends on MSs channel gain as well as their spatial distribution within a BS coverage area. Finally, in step 5, we determine $T_{i+1}$ for next iteration by eliminating all MSs whose channel vectors are not semiorthogonal to $\pi_i$ within a certain threshold $\alpha$. If $T_{i+1}$ is not an empty set and $|\delta| < M^{opt}$, then go to step 2. Otherwise, the algorithm terminates and returns a set of scheduled MSs.

Note that Algorithm 1 requires two thresholds $\Omega$ and $\alpha$. In practice, one can try multiple thresholds in order to select/determine the one that minimizes the number of total transmissions. We now propose our second technique.

Algorithm 2 is described as follows: in step 1, $V$ represents the set of all $K$ MSs and $Q \times M$ represents the number of all possible beams in a codebook. $V_i$ is the set of all assigned MSs in beam $i$ and is initially empty where $i \in \{1, 2, \ldots, Q \times M\}$. In step 2, we select beam $i^*$ for MS $z_j$ whose SINR is maximized (the best assigned beam) over all possible beams. In step 3, we update the value of $V$ and $V_i$. Finally, in step 4, if $V$ is not an empty set, then go to step 2. Otherwise, the algorithm terminates, and returns a set of MSs in $V_1, V_2, \ldots, V_{Q \times M}$. If some set of MSs (e.g., $V_i$ for beam $i$) is an empty set, this implies that the BS will not transmit any packets over that beam whose power allocation is zero. Note that Algorithm 2 minimizes the variance of SINRs of all MSs in each beam.

5.2 Technique 2: Codebook-Based Scheduler. The second technique, as described in the pseudocode shown below, uses two algorithms: Algorithm 2, which determines the best beam for each MS whose SINR is maximized and Algorithm 3, which further minimizes the number of transmissions. The outcome will be a set of MSs in $V_1, V_2, \ldots, V_{Q \times M}$, an optimized codebook $L$, and the expected number of transmissions $T_Q$. 

Pseudocode for technique 2

```
begin
    $V = \{z_1, z_2, \ldots, z_K\}$
    $V_1 = V_2 = \cdots = V_{Q \times M} = \emptyset$ (empty set)

    Run Algorithm 2
    {determine the best beam for each MS}
    Run Algorithm 3
    {further minimize the number of transmissions}

    return a set of MSs in $V_1, V_2, \ldots, V_{Q \times M}$, an optimized codebook $L$, and the expected number of transmissions $T_Q$.

Algorithm 2 is described as follows: in step 1, $V$ represents the set of all $K$ MSs and $Q \times M$ represents the number of all possible beams in a codebook. $V_i$ is the set of all assigned MSs in beam $i$ and is initially empty where $i \in \{1, 2, \ldots, Q \times M\}$. In step 2, we select beam $i^*$ for MS $z_j$ whose SINR is maximized (the best assigned beam) over all possible beams. In step 3, we update the value of $V$ and $V_i$. Finally, in step 4, if $V$ is not an empty set, then go to step 2. Otherwise, the algorithm terminates, and returns a set of MSs in $V_1, V_2, \ldots, V_{Q \times M}$. If some set of MSs (e.g., $V_i$ for beam $i$) is an empty set, this implies that the BS will not transmit any packets over that beam whose power allocation is zero. Note that Algorithm 2 minimizes the variance of SINRs of all MSs in each beam.
Using all Q sets, a codebook may not be the best solution. In other words, using a large number of sets of a codebook implies that the BS spends an amount of time to successfully retransmit all lost packets while all MSs most likely have high SINR levels. However, when we lower SINR values for a certain set of MSs by eliminating some sets of a codebook and assigning these MSs to the remaining sets, the overall number of transmissions can be reduced. Therefore, we propose an algorithm to further minimize the number of transmissions as shown in Algorithm 3.

Algorithm 3 works as follows: in step 1, we have the original sets of a codebook \( L = \{l_1, l_2, \ldots, l_Q\} \) and the groups of MSs in each \( V_1, V_2, \ldots, V_{Q \times M} \). Then, we compute the expected number of transmissions \( T_Q \) where \( Q \) is the number of sets of a codebook. In step 2, we determine the maximum PER over all MSs assigned to each set \( l_q \) of a codebook denoted by \( X_q \). The best set of a codebook \( \left(l^*_q\right) \), that is, the set that has the lowest value of the maximum PERs among all sets of a codebook, is determined in step 3. Because the set with the highest value of maximum PERs dominates the transmission performance in the system, when we eliminate the set \( l^*_q \), the number of transmissions is likely to decrease. Then, we assign all MSs in the set \( l^*_q \) to all other existing sets as done in step 4. In step 5, we compute the expected number of transmissions \( T_{Q-1} \) after eliminating the set \( l^*_q \). In step 6, if \( T_{Q-1} \leq T_Q \), then we go to step 2 for further codebook optimization. Otherwise, the algorithm returns the optimized codebook together with a set of MSs in \( V_1, V_2, \ldots, V_{Q \times M} \) and the expected number of transmissions \( T_Q \) as described in step 7.

6. Performance Evaluation

We consider a downlink channel consisting of one BS and multiple MSs. The BS and each MS are, respectively,
Data transmission phase

Retransmission phase

![Broadcast transmission/retransmission schemes.](image)

**Figure 6:** Broadcast transmission/retransmission schemes.

- BS (a) Uniform distribution
- BS (b) 3-cluster distribution
- BS (c) Random distribution

**Figure 7:** Simulated network topology scenarios with different MS distributions.

For the purpose of evaluating the performance of our proposed scheduling algorithms, we consider evaluating four retransmission schemes: 1-Omni ARQ (i.e., traditional retransmission technique), 1-Omni NC (i.e., traditional with NC only), BF-ARQ (i.e., BF with traditional retransmission technique), BF-NC with SINR (proposed technique 1: greedy-selection-based scheduler), and BF-NC with codebook (proposed technique 2: codebook-based scheduler). Both 1-Omni ARQ- and 1-Omni NC-based retransmission schemes use one transmit antenna element, and retransmissions of lost packets are respectively done with ARQ and NC techniques only. BF-ARQ-based retransmission scheme is to use BF to retransmit lost packets to a group of MSs with ARQ techniques. The group formation for BF-ARQ is based on technique 1 (Algorithm 1). The proposed BF-NC-based retransmission schemes use either technique 1 (Algorithm 1) or technique 2 (Algorithms 2 and 3). A graphical illustration of these schemes is given in Figure 6.

6.1. Uniform MS Distribution on a Circle. First, we study the broadcast scenario shown in Figure 7(a), where there are 120 uniformly distributed MSs, all having the same channel gain. In broadcast scenarios, the BS transmits a packet to all MSs at once. One example of broadcast transmissions is digital TV broadcasting, where all MSs watch the same TV channel at the same time. Therefore, the BS needs only to transmit data packets omnidirectionally, which makes them reach all MSs. We assume that the BS uses one antenna to do so, and there is no need for applying BF techniques during data transmissions.

Recall that the BS uses one transmit antenna during the data transmission phase, making signals/packets (transmitted omnidirectionally) reach all MSs. In Figure 8(a), we show the average loss rate of packet transmissions during these omnidirectional transmissions as a function of the transmit power. Figures 8(b) and 8(c) show the total number of transmitted packets and throughput gain, respectively, as a function of power for all simulated retransmission schemes.

**Figure 8(b)** shows that 1-Omni NC-based retransmission scheme performs the best in this network scenario when compared with all the other retransmission schemes, followed by the BF-NC with codebook, the BF-NC with SINR, the 1-Omni ARQ, and then the BF-ARQ. The achieved performances of the 1-Omni NC scheme are expected in this network topology, since all MSs experience the same channel gain and signal strength. We will see later that this is no longer the case when considering more realistic scenarios where MSs may experience different gains. For the BF-NC with SINR (i.e., with greedy selection) scheme, SINR thresholds are set to 3 levels (6.75, 7.25, and 8.00) for grouping MSs in each beam. For completeness and to assess the impact of incorporating the NC technique, we also measure and plot in Figure 8(c) the throughput gains of the 1-Omni NC, the BF-ARQ, the BF-NC with SINR,
and the BF-NC with codebook over the 1-Omni ARQ-based retransmission scheme. In all figures, a throughput gain of ratio 1 means that both retransmission schemes use the same number of total transmissions. The figure shows that the 1-Omni NC- and the BF-NC-based retransmission schemes reach throughput gains of ratio 2.2 (i.e., 120% gain) and 1.5 (i.e., 50% gain), respectively. Even though both of the proposed BF-NC-based retransmission schemes have similar results, we lean towards the codebook-based one since it does not require multiple levels of SINR thresholds while still achieving similar performances.

6.2. Clustered MS Distribution on a Circle. In this section, we consider studying a network topology with a clustered MS distribution. Specifically, we consider the network topology scenario of Figure 7(b), where 60 MSs are grouped into 3 clusters each having 20 MSs. All MSs in each cluster are close to one another, and the channel gains for all MSs belonging to the same cluster are the same (or close). Figure 9(a) shows the average loss rate of omnidirectional packet transmissions (without including retransmissions) as function of the transmit power. In Figure 9(b), we show the number of
total transmissions (including retransmissions) under each of the studied retransmission schemes as function of the transmit power. First, observe that unlike the case when MSs are uniformly distributed (as illustrated in Section 6.1), when MSs are not uniformly distributed, the proposed retransmission schemes (both BF-NC with SINR and BF-NC with codebook) outperform not only the traditional retransmission scheme (i.e., 1-Omni ARQ and BF-ARQ), but also the traditional scheme with NC (i.e., 1-Omni NC scheme). The total number of transmissions needed under the proposed schemes is substantially lesser than that needed under the 1-Omni ARQ and the 1-Omni NC, especially for low transmit power levels. This is because the BS can now efficiently allocate the power to a set of active BF beams instead of spreading the power in all directions, strengthening then the signals received at intended MSs, and as a consequence, increasing MSs SINRs.

The second observation that we can draw from the figure is that both proposed techniques achieve very similar performances in terms of the total number of transmissions. This is because when all MSs in the same cluster have similar channel gain, each coded packet most likely comes from all MSs whose packet loss rates are similar. In other words, this maximizes the number of lost packets to be coded at once.

We now plot the throughput gains of the proposed schemes and the 1-Omni NC scheme when compared with
the traditional scheme (i.e., 1-Omni ARQ). As shown in Figure 9(c), the throughput gains for the 1-Omni NC- and the BF-NC-based retransmission schemes can reach up to 120% (ratio of 2.2) and 175% (ratio of 2.75), respectively. Note that the incorporation of NC into the 1-Omni scheme results in almost halving the number of total transmissions, whereas the incorporation of these same NC techniques into BF-based schemes (whether SINR or codebook based) reduces the number of total transmissions by almost a third.

6.3. Random MS Distribution inside the Circle. Finally, we study a network topology with a random MS distribution. In particular, we consider the network topology scenario of Figure 7(c), where 120 MSs are randomly distributed at distances from the base station such that their channel gains are uniformly distributed between 0.7 and 1.3. Furthermore, the angles are normally distributed with the mean of 90 and standard deviation of 30 degrees. Any station located outside the range of [0, 180] degrees ignored.

The average loss rate of packet transmissions during these omnidirectional transmissions is shown in Figure 10(a). The performance of BF-based retransmission schemes in this case is under that of the 1-Omni NC scheme as in Figure 10(b). Specifically, the 1-Omni NC scheme has
the best result followed by NC-based retransmissions (i.e., BF-NC with SINR or codebook), the 1-Omni ARQ, and the BF-ARQ, respectively. Finally, Figure 10(c) shows the throughput gain with respect to the 1-Omni ARQ scheme. This is intuitively plausible as when mobile stations are dispersed too much, beamforming becomes less effective. In this case, omnitransmission makes sense and NC provides additional benefit with fewer retransmissions. Thus, 1-Omni NC transmission scheme performs the best.

6.4. Impact of Number of MSs. In this section, we want to study the impact of the number of MSs on the performance behaviors. In this simulation, we use the network topology of Figure 7(a), where the number of MSs is varied from 120 to 300. In all scenarios, all MSs have the same channel gains. Figure 11(a) shows how the number of transmissions changes as the number of MSs is varied for each of the proposed retransmission schemes. As expected, when increasing the number of MSs, the BS tends to spend more time to successfully retransmit all lost packets to all corresponding MSs. The number of transmissions when using the BF-NC with codebook retransmission scheme is less than that of the BF-NC with greedy selection (with SINR) retransmission scheme. Note that the number of MSs affects the throughput gain of both BF-NC-based retransmission schemes; this is shown in Figure 11(b). The higher the number of MSs, the higher the chances of combining more lost packets before retransmitting, thus leading to higher throughput gains (lesser transmissions).

7. Conclusion

This paper proposes two joint BF/NC-based scheduling algorithms for wireless broadcast downlink channels. The first algorithm relies on greedy selection of MSs based on their SINRs, whereas the second algorithm uses codebook to group and schedule MSs. The greedy selection approach consists of using MSs orthogonality and their physical locations to maximize their signal quality over a wireless network. The codebook-based approach, on the other hand, uses a predefined codebook to generate multiple sets of orthogonal beams.

Our results show that both of the proposed techniques/algorithms increase the overall throughput by minimizing the number of transmission attempts.

Our simulations indicate that our techniques reduce the number of retransmission attempts significantly, resulting in achieving high throughput gain of up to 175% of broadcast transmissions when compared to that of traditional ARQ retransmission approaches.

Appendix

In this section, we prove that our BF-NC scheduling problem is NP-hard. First of all, we know that the well-known problem “scheduling unrelated parallel machines (SUPM)” is NP-hard [19]. To prove that our BF-NC-based scheduling problem is NP-hard, we need to show the reduction from the SUPM problem to our problem. In the SUPM problem, there is a set of $K$ jobs denoted by $J = \{p_1, p_2, \ldots, p_K\}$ and a set of $M$ parallel machines with their related speed...
factors denoted by \( S = \{s_1, s_2, \ldots, s_M\} \). For generalization of identical machines, they can run at different speeds but do so uniformly. For each machine \( i \) with a speed factor \( s_i \), the processing time for job \( j \) (\( p_j \)) on machine \( i \) denoted by \( p_{ij} \) can be computed as \( p_{ij} = p_j / s_i \). Note that each machine \( i \) can process only one job at a time. Therefore, the total time on machine \( i \) is \( \sum p_{ij} \). The goal is to find a scheduler that minimizes the makespan such that

\[
\begin{align*}
\text{minimize} & \quad \max_i \left\{ \sum_j x(i, j) p_{ij} \right\}, \\
\text{s.t.} & \quad \sum_i x(i, j) = 1, \\
& \quad x(i, j) \in \{0, 1\},
\end{align*}
\tag{A.1}
\]

where \( i \in \{1, 2, \ldots, M\} \) and \( j \in \{1, 2, \ldots, K\} \). Formally, an instance of an SUPM problem is a 3-tuple problem as follows:

\[
\langle \{p_1, p_2, \ldots, p_K\}, \{s_1, s_2, \ldots, s_M\}, \{x(1, 1), \ldots, x(M, K)\} \rangle.
\tag{A.2}
\]

Given an instance of an SUPM problem, we can construct the following version of our BF-NC-based scheduling problem by reduction as follows. All MSs are in a set \( V \). A set of lost packets belonging to MS \( z_k \) is \( p_k \). In particular, the job \( p_k \) in original problem is reducible to our problem as a set of lost packets \( n_k \) for MS \( z_k \) (i.e., \( n_k = p_k \)). For each beam \( i \), we randomly assign MS \( z_k \) with a set of lost packets \( p_k \) to a group of MSs \( V_i \) over all \( M \) groups as shown in Figure 12, where \( V = \bigcup_V V_i \) and \( i \in \{1, 2, \ldots, M\} \). In this case, we set \( x(i, k) = 1 \) for assigning a set of lost packets \( p_k \) to a group \( V_i \) (for beam \( i \)). Note that this reduction is the special case of our BF-NC-based scheduling problem when all lost packets in each group cannot be combined. Therefore, the reduction for variables \( p_k \) to \( n_k \) can be done in \( O(1) \).

Suppose we have the optimal solution to solve a BF-NC-based scheduling problem. To be complete, we can show that the optimal solution to our BF-NC-based scheduling problem is the optimal solution to an SUPM problem as follows. First, we use only one set of orthogonal BF beams from a given codebook. For example, the BS uses only the \( q \)th set, then we set \( l_q = 1 \) and \( l_p = 0 \) in (8) where \( p \neq q \). Therefore, we have

\[
\begin{align*}
\text{minimize} & \quad \max_i \left\{ \frac{N_{V_i}}{1 - \max_j x(i, j) \text{PER}_{ij}} \right\}, \\
\text{s.t.} & \quad \sum_i x(i, j) = 1, \\
& \quad x(i, j) \in \{0, 1\},
\end{align*}
\tag{A.3}
\]

where \( N_{V_i} \) and \( \text{PER}_{ij} \) are a set of coded/uncoded lost packets based on an NC technique in beam \( i \) and packet error rate for MS \( z_j \) in beam \( i \), respectively. Recall that all lost packets in each beam \( i \) cannot be combined in this special case. Thus, a set of lost packets \( n_j \) for MS \( z_j \) strictly depends on \( \text{PER}_{ij} \), where \( n_j \subseteq N_{V_i} \). In such case, we can rewrite (A.3) as follows:

\[
\begin{align*}
\text{minimize} & \quad \max_i \left\{ \sum_j x(i, j) \frac{n_j}{1 - \text{PER}_{ij}} \right\}, \\
\text{s.t.} & \quad \sum_i x(i, j) = 1, \\
& \quad x(i, j) \in \{0, 1\}.
\end{align*}
\tag{A.4}
\]

In (A.4), the time to successfully retransmit all lost packets in \( N_{V_i} \) is the summation of the time to retransmit all \( n_j \) where \( n_j \subseteq N_{V_i} \).

Comparing (A.1) to (A.4) and assuming there exists a function for reduction from \( p_{ij} = p_j / s_i \) to \( n_j / (1 - \text{PER}_{ij}) \), all instances of an SUPM problem are correctly reducible to our BF-NC-based scheduling problem. That is

\[
\begin{align*}
\text{minimize} & \quad \max_i \left\{ \sum_j x(i, j) \frac{p_j}{s_i} \right\}, \\
\iff \quad \text{minimize} & \quad \max_i \left\{ \sum_j x(i, j) \frac{n_j}{1 - \text{PER}_{ij}} \right\}.
\end{align*}
\tag{A.5}
\]

The solution to (A.4) is a set of \( x(i, j) \) for all \( i \) and \( j \). From (A.5), both objective functions can be reducible to each other and vice versa. That means the optimal solution to our BF-NC-based scheduling problem in (A.4) is the optimal solution to an SUPM problem in (A.1). Furthermore, the SUPM problem can be reducible to our problem in polynomial time. This concludes that our problem is NP-hard.

References


