This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2011 proceedings.

YMMV: Multiple Session Multicast with MIMO

Hong Xu, Jin Jin, Baochun Li
Department of Electrical and Computer Engineering
University of Toronto
{*henryxu, jinjin, bli*}*@eecg.toronto.edu*

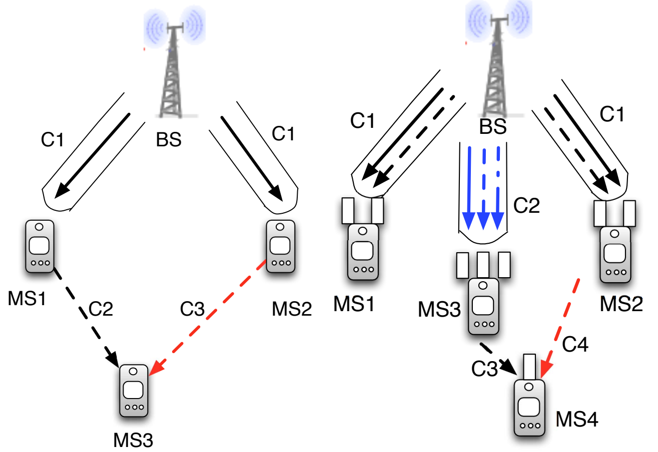
***Abstract*—Multicast is an important application in cellular networks. The 4G technologies, including WiMAX and LTE, invariably adopt Multiple-Input-Multiple-Output (MIMO) to facilitate spatial multiplexing and fundamentally increase channel capacity. However, state-of-the-art multicast protocols are designed to perform in single-hop mode with a single session, leading to under-utilization of the scarce spectrum resource.**

In this paper, we propose YMMV, a novel multicast protocol that jointly considers MIMO and cooperative communications in OFDMA networks. The base station transmits data in multiple *sessions* using multiple antennas on the same channel to exploit spatial multiplexing in MIMO. Further, cooperative transmission on different channels among users is also utilized. We tackle the resulted session scheduling problem in YMMV, where the multi-channel characteristic of OFDMA further aggravates the difficulty of efficient algorithm design. With rigorous analysis and extensive simulations, we show that our multi-session multicast protocol is able to improve throughput performance significantly.

I. INTRODUCTION

Multicast serves as an important multimedia application in cellular networks, for which transmission scheduling plays a critical role. Existing multicast protocols work in a single-hop single-channel fashion [1]–[3]. The base station (BS) uses the most robust modulation and coding scheme to provide reliable transmissions to all users. Essentially, the multicast rate is limited to the lowest possible rate among all users, which under-utilizes the wireless spectrum by a substantial margin. The state-of-the-art research in multicast scheduling schemes takes advantage of multi-hop multi-channel communication [4], [5] in OFDMA networks [6]. For example, as shown in Fig. 1(a), the BS uses a high rate to multicast data on channel 1 (C1), and mobile stations (MSs) with good channel qualities (MS 1 and 2) can receive the data and cooperatively help the ones in poor channel conditions (MS 3) using orthogonal channels (C2 and C3).

Such multi-path transmissions are able to boost throughput. However, transmission on each hop is still hobbled by the holdover from traditional cellular communication — the insistence on transmitting a single *session* of data on each channel. The adoption of Multiple-Input-Multiple-Output (MIMO) [7] technology by 4G networks as well as advanced devices such as the iPhone makes the use of multi-session communications through multiple antennas realistic. As shown in Fig. 1(b), for a given channel, the transmitter and receiver can communicate multiple independent sessions as long as both of them are equipped with multiple antennas. Such spatial multiplexing in MIMO changes the fundamental relationship between power and capacity, and enables linear scaling of transmission rate in theory [8].



(a) Conventional multicast. (b) Multi-session multicast with MIMO.

Fig. 1. Illustrative examples of a multi-session multicast framework supported by MIMO in OFDMA networks.

In this paper, we investigate multicast scheduling with multiple sessions, multiple hops and multiple channels, which is barely explored in the literature. The intuitive idea is shown in Fig. 1(b). The BS may multicast data in two sessions using MIMO simultaneously through channel C1, leading to a higher multicast rate than single-session transmissions. MS 1 and 2 have two antennas each, and MS 4 has three antennas. All of them can correctly receive two independent sessions of data, which roughly doubles the throughput. Even better, the BS can communicate with MS 3 three independent sessions simultaneously through an additional channel (C2). In addition to single-hop transmissions, users are allowed to cooperatively help each other on orthogonal channels. MS 2 and 3 can both forward data to MS 4 using separate channels, who only has one antenna and can not correctly decode the multicast data issued in multiple sessions.

Therefore, with the same number of channels, several independent sessions of transmissions can be performed simultaneously instead of transmitting one session on one channel. Channels are better utilized with MIMO multiplexing, and throughput performance can be largely improved.

Several technical challenges arise from the introduction of MIMO to multicast. Specifically, now we need to schedule the multicast transmissions in multiple sessions according to

978-1-4244-9268-8/11/$26.00 ©2011 IEEE

channel conditions and user antenna equipment, which clearly adds to the complexity of developing efficient algorithm to solve it. Moreover, we also need to schedule the cooperative transmissions among users optimally in order to further improve multicast throughput.

The highlight of our contribution in this paper is a cross-layer optimization framework for transmission scheduling in multicasting multiple sessions with MIMO, which we call *YMMV*. The name coincides with the common acronym of "your mileage may vary," which also reflects the unique facet of our multicast protocol that different users may enjoy different multicast rates depending on the number of sessions they receive, which is essentially related to their number of antennas and channel conditions.

The remainder of the paper is organized as follows. In Sec. II, we review the basics of MIMO and related work on multicast scheduling. In Sec. III, the YMMV optimization framework is presented and the session scheduling problem is solved. We present simulation results to evaluate the performance of our algorithms along with the analysis in Sec. III. Corroborating our intuition, multicast performance is substantially improved with our design. We conclude our paper in Sec. IV.

II. BACKGROUND

A. MIMO Basics

MIMO serves as the cornerstone of our protocol, and is instrumental towards most of its advantages. MIMO has two basic working scenarios. *Spatial diversity* improves the reliability and range of transmission by sending and/or receiving redundant streams of information in parallel along different spatial paths between transmitter and receiver antennas. The real excitement around MIMO is that the independent paths between multiple antennas can be used to much greater effect than simply for diversity to boost SNR. *Spatial multiplexing* takes advantage of this extra degree of freedom to send *independent* streams of information, called *sessions* in this paper, at the same time over the same frequency. Conceptually, the received signals can be seen as a set of linear equations with channel gains being the coefficients, and the solutions of this linear system correspond to the transmitted signals [8].

In practice, channels may achieve linear gains in capacity with the use of spatial multiplexing. We can roughly express the capacity of multi-session MIMO as $BN\log_2(1+\rho)$, where B is the bandwidth, N is the number of antennas, and ρ is the SNR [8]. This is valid under the constraint that the number of sessions is no larger than the minimum number of antennas on the transmitter and receiver. In general, the number of antennas at the BS is larger than that of any MS. Therefore, the linear capacity gain of using multiplexing MIMO can be assumed.

B. Related Work

Optimization in multi-channel networks has drawn a substantial amount of attention in recent years, especially in wireless mesh networks and OFDMA networks. [9] and [10] formulated mathematical models for multi-channel multi-radio networks and solveed the joint multi-commodity routing and channel assignment problems. Cooperative and relay communications have also been considered as another dimension of optimization. [11] proposed solutions for a joint optimization of channel assignment, relay strategy selection and power allocation in OFDMA cellular networks based on conventional Amplified-and-Forward and Decode-and-Forward schemes. [12] considered network coding assisted cooperative diversity in OFDMA cellular networks. [13] studied opportunistic scheduling in WiMAX relay networks. These works considered multiple unicast sessions of independent data, while our work considers multicast of the same content to a group of users.

Multicast in multi-channel networks is less well studied. Zeng *et al.* [4] tackled the channel assignment problem for multicasting in mesh networks, and Gopinathan *et al.* [14] extended it to multi-radio mesh networks. One recent work [15] studies the interactions of MIMO with higher layer protocols. Cooperative communications were modeled in multicast systems in our previous work [5] for WiMAX. These works only considered single-session multicast, and interference on the same channel is treated as noise, i.e. MIMO is not modeled. Our work, in contrast, is an early attempt to utilize MIMO in multicast services in OFDMA cellular networks.

III. A SCHEDULING FRAMEWORK FOR MULTI-SESSION MULTICAST WITH MIMO

In this section, we present our YMMV optimization framework with MIMO. We start by introducing the system model. The cellular network operates in a time-slotted fashion, where the BS serves as the multicast sender and MSs (also referred to as nodes) as the receivers. We assume quasi-stationary channel conditions, i.e. they remain stationary during one period of time, and vary independently from one time slot to another. The BS is equipped with multiple antennas, and MSs are equipped with one or multiple antennas respectively. The number of antennas at the BS is larger than that of any MS. Therefore, one BS can serve multiple MSs at the same time, and MIMO communications can be performed whenever possible.

All nodes work in full-duplex mode, and concurrent communication with multiple nodes in both downlink and uplink is possible on different channels. In order to show the benefits provided by MIMO multiplexing in multicast, here we assume that the number of available orthogonal channels for cooperative communications is greater than needed. We will study more complicated cases with limited channel resources in the future.

A. Optimization Framework

Our multicast scheduling problem with MIMO can be formulated as an optimization outlined as follows. The objective can be stated as to find the optimal session allocation for MIMO and cooperative communications to maximize the aggregate throughput of all users, under the commonly used proportional fairness criterion [16]. Mathematically,

$$\max \quad \sum_{i\in\zeta} \frac{U_i(t)}{\overline{r}_i(t)} \tag{1}$$

where $U_i(t)$ is the throughput of node i at time slot t, taking into account transmissions from both the BS and cooperating nodes. ζ is the set of nodes in the multicast group, and its cardinality is G. $\overline{r}_i(t)$ denotes the average throughput of node i over the time horizon $[1, t]$. It serves as a time-varying weight with which proportional fairness can be achieved in the long run [16].

$U_i(t)$ can be calculated as:

$$U_i(t) = \sum_{s=1}^{N} L_s(t) X_i^s(t) R_i(t) + \sum_{g\in\zeta} R_{gi}(t). \tag{2}$$

We assume that the content to be multicast can be dynamically divided and coded into MIMO sessions according to channel and session allocation to enable adaptive transmission in the wireless medium [17]. The spectrum is divided into a number of OFDMA subchannels. The BS uses a certain number of subchannels to multicast the basic layer of the content with the most robust rate to provide reliable transmission for all subscribers. Meanwhile, it uses other channels for MIMO communication to send additional layers that contain finer details of the content. As such, we do not consider the complementary challenges of producing these independent sessions, and assume that the higher throughput a node enjoys, the better the perceived quality-of-service in the multicast service. N is the total number of channels that the BS can use for direct multicast. Thus, $s \in \{1, \cdots, N\}$.

As we elaborated in Sec. I, the BS multicasts data using MIMO whenever possible. On each channel, data is transmitted in one or several independent *sessions*. To model MIMO without exposing details in the PHY layer, $L_s(t)$ is used to denote the number of sessions for MIMO transmissions on channel s in (2). It essentially represents our design of utilizing MIMO in OFDMA networks.

$X_i^s(t)$ represents the actual number of sessions that node i can correctly decode when the BS multicasts data in $L_s(t)$ sessions on channel s. Apparently, when $L_s(t)$ is larger than the number of antennas at MS i, decoding is not possible and all the data on this channel will be discarded by i. Otherwise, multi-session data can be correctly received. Thus we have:

$$X_i^s(t) = \begin{cases} 1, & \text{if } L_s(t) \le a_i \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

where integer $a_i \in [1, S]$ is the number of antennas of MS i. $L_s(t)$ is clearly upper bounded by the number of antennas at the BS, a_0.

$$L_s(t) \le a_0. \tag{4}$$

$R_i(t)$ is the per-session throughput from the BS to node i on one channel. Here we do not consider channel diversity for simplicity. Then, we have:

$$0 \le R_i(t) \le C_i(t) = B \log_2(1 + \frac{P_{BS} g_i}{N_i}) \tag{5}$$

where $C_i(t)$ is the Shannon capacity of this session. We assume the BS uses equal power P_{BS} on each channel, and the noise power N_i is independent across nodes. The channel gain g_i is independent as well and models the user diversity.

Now we can see that the first term $\sum_{s=1}^{N} L_s(t) X_i^s(t) R_i(t)$ from Eq. (2) represents the throughput of direct transmissions from the BS to all nodes. The non-trivial session allocation clearly has a major impact on throughput. If we increase $L_s(t)$, nodes with fewer than $L_s(t)$ antennas will fail to receive any data on channel s, and suffer from throughput degradation. On the contrary, nodes with no fewer than $L_s(t)$ antennas will enjoy throughput improvement, since they receive more sessions from the BS.

The second term in Eq. (2) shows that nodes with slow downloading can be compensated by cooperation. $R_{gi}(t)$ is the cooperative transmission rate achievable from node g to i. Clearly,

$$0 \le R_{gi}(t) \le C_{gi}(t) \tag{6}$$

which means that the cooperative transmission rate is bounded by the capacity on the link, denoted as $C_{gi}(t)$. It is challenging to schedule transmissions in a cooperative fashion. Relays do not have knowledge about which packets their neighbors need. Blindly "pushing" packets that are not needed to other peers will incur a substantial degree of overhead. To address this challenge, we propose to take advantage of the favorable rateless properties of *network coding* [18]. With this technique, all packets are encoded with random linear codes, and all coded data blocks could be considered equally useful and innovative. As random network coding is employed, a packet is innovative if it is linearly independent from the other packets from the same segment, which is satisfied with high probability when the field size is reasonably large [19].

Since the data is fully mixed, relays can freely "push" innovative blocks to their downlink multicast members. Overhead can be substantially mitigated in cooperative communication. With this design, the cooperative transmission rate is limited by the amount of innovative data that node g is able to contribute to node i. Here we dictate that only a node with comparatively more received data can help another with less data in order to achieve better fairness among the nodes of the multicast group. We further constrain the amount of data exchanged between a pair of cooperating nodes by the difference in their buffers. This is expressed in (7), where $B_g(t)$ denotes the amount of data buffered at node g at time slot t, and $B_i(t)$ indicates the same information at node i. T is the duration of one time slot. It is easy to get from this constraint that $R_{gi}(t) = 0$ if $g = i$.

$$R_{gi}(t) \le \max\{0, \frac{B_g(t) - B_i(t)}{T}\} \tag{7}$$

Besides the aforementioned constraints, the system throughput $U_i(t)$ is also constrained as the total amount of data that

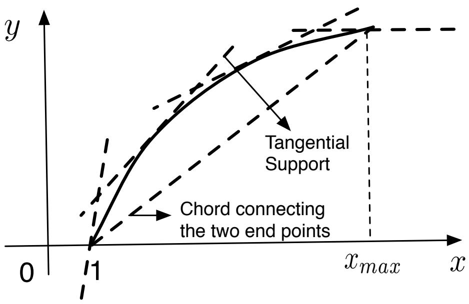


Fig. 2. Intuition on linearization for logarithmic relationship, which uses a four-point tangential approximation.

each node receives can not exceed the amount the BS is able to provide:

$$\begin{aligned} U_i(t) \leq \sum_{h=1}^{t}\sum_{s=1}^{S} L_s(h)R_i(h) - \frac{B_i(t)}{T} \quad &\Rightarrow \\ \sum_{g\in\zeta} R_{gi}(t) \leq \sum_{h=1}^{t-1}\sum_{s=1}^{S} L_s(h)R_i(h) - \frac{B_i(t)}{T} &+ \\ \sum_{s=1}^{S} L_s(t)R_i(t)(1 - X_i^s(t)). \end{aligned} \tag{8}$$

Overall, session allocation in the basic YMMV framework can be formulated as an optimization problem over $L_s(t)$, with objective (1), subject to constraints (2)–(8).

B. Optimal Solution

The formulated problem is a non-linear integer programming (NIP) problem, which is NP-hard in general. In the following, we discuss a *linearization* technique and use the *branch-and-bound* algorithm to solve the resulted LP with polynomial-time complexity.

We first relax the integer variables $L_s(t)$ into fractional values in $[0, a_0]$. In addition, we need to linearize constraint (5) which is not convex. To address this challenge, we adopt the Reformulation-Linearization Technique (RLT) [20] that produces an LP relaxation for an underlying nonlinear problem by providing a tight upper bound. According to RLT, we linearize the logarithmic relationship in (5) using polyhedral outer approximation with several tangential supports [20]. The intuition is shown in Fig. 2.

Readers may observe that constraints (3) and (7) contain nonlinear relationships as well. However, they do not actually generate non-linear constraints in the optimization. The buffer information $B_i(t)$ can be captured by each node, and this information can be reported to the BS through message exchange. $X_i^s(t)$ can be explicitly expressed in constraint (2), as the number of antennas for each node a_i is global knowledge in the system.

With linear relaxations we can now apply the *branch-and-bound* algorithm to the resulted LP. With this approach, we aim to provide a $(1-\epsilon)$-optimal solution, where ϵ is a small positive constant reflecting our desired optimality gap. We first solve the LP and get fractional solutions $\hat{L}_s(t)$ and the corresponding upper bound (UB) of the objective. Over the fractional solutions, we then conduct a local search to find a feasible lower bound (LB) of the objective. In our problem, we adopt randomized rounding on $\hat{L}_s(t)$ to its closest integer to get LB, while ensuring the solution feasibility.

If $LB \geq (1-\epsilon)UB$, then we have obtained the desired $(1-\epsilon)$-optimal solution. If not, we have to close the gap through a tighter linear relaxation. This could be achieved by selecting the optimizing variable with maximum relaxation error, and dividing its value set into two by its value in the relaxation solution. In our problem, we choose a $\hat{L}_s(t)$ with a maximum value of relaxation error captured by $\min\{\lfloor\hat{L}_s(t)\rfloor, \lceil\hat{L}_s(t)\rceil\}$, and divide the original problem into two subproblems with $\hat{L}_s(t)$ equal to $\lfloor\hat{L}_s(t)\rfloor$ and $\lceil\hat{L}_s(t)\rceil$, respectively.

For the two subproblems, we again solve the LP relaxation and run local search to get their bounds: (UB_2, LB_2) and (UB_3, LB_3). We update $UB = \max\{UB_2, UB_3\}$ and $LB = \max\{LB_2, LB_3\}$. Then, if $LB \geq (1-\epsilon)UB$, the algorithm is terminated. Otherwise, we will iteratively repeat the entire procedure until it is so. During this process, we remove any subproblem i when $(1-\epsilon)UB_i \leq LB_i$. It has been shown that under general conditions, a branch-and-bound procedure always converges efficiently [21].

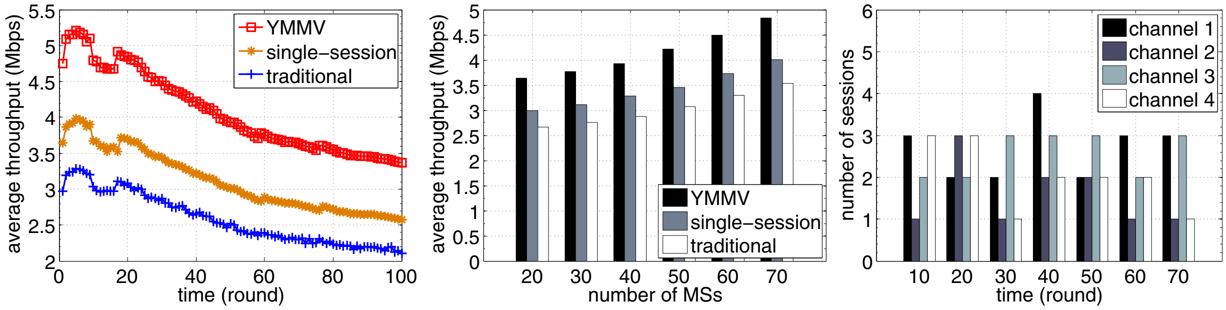
C. Performance Evaluation

We now resort to extensive simulations to evaluate the performance of our YMMV protocol. To be realistic, simulations are performed by emulating the multicast broadcast service (MBS) in WiMAX with typical parameters according to the IEEE 802.16 standard [22] and the WiMAX system evaluation methodology released by the WiMAX forum. The BS multicasts a large file to all MSs. Each MS is allowed to move randomly in the service area, with a randomly chosen initial location. Multi-path Rayleigh fading is simulated since user mobility is present.

To effectively capture the benefits provided by MIMO, we set the number of antennas each MS has to be uniformly random from 1 to 5. We also assume that all the transmissions on different paths of MIMO are independent. Further, there are a total of 20 channels with equal bandwidth available for the BS multicasting data.

To evaluate performance, we compare three multicast protocols: our YMMV framework, referred to as "YMMV", cooperative multicast without MIMO, referred to as "single-session;" and the traditional multicast without MIMO and OFDMA, referred to as "traditional."

Fig. 3(a) shows the average throughput across MSs for all three protocols. We observe from the results that "YMMV" performs best, with 20% gain compared to "single-session." It further outperforms "traditional" by a larger margin of 35%. This coincides with our intuition that multi-session multicast supported by MIMO fits the design of 4G networks well, and is able to achieve significant throughput improvement due to its effective use of the wireless spectrum. A trend to notice



(a) Throughput with 20 MSs

(b) Average throughput vs. No. of MSs

(c) Session allocation for channels.

Fig. 3. Throughput performance of three multicast protocols in a realistic WiMAX MBS scenario.

is that the average throughput is slowly decreasing over time. The reason is that our objective takes fairness into account, which makes the optimization favor "slower" MSs over time.

Another interesting result we observed is that the margin that "YMMV" outperforms "single-session" and "traditional" becomes more substantial with an increasing number of MSs, as in Fig. 3(b). This observation indicates that more MSs create a higher degree of antenna diversity and cooperation, leading to higher throughput.

Fig. 3(c) further explores the advantages of multi-session multicast and shows the results of session allocation on a randomly selected subset of 4 channels. We can see that the BS uses different number of sessions to multicast data on different channels. As time goes and MSs move around the service area, the BS dynamically tunes this allocation to achieve maximum resource utilization. From the results, we also observe that the distribution of the number of sessions allocated to multiple channels is clustered around $[2, 3]$. It clearly shows that fairness is considered in the protocol, as not every node is equipped with multiple antennas.

IV. CONCLUDING REMARKS

In this paper, we propose YMMV, a novel multi-session multicast optimization framework with MIMO in OFDMA networks. MIMO provides abundant opportunities for spatial multiplexing. Our YMMV framework also exploits cooperative diversity, leading to significant capacity improvement for multicast services. Tightly integrated with 4G cellular networks, our work represents an early attempt to address the session scheduling issue with the use of MIMO and cooperative communications. Throughput improvement is demonstrated compared to conventional approaches.

REFERENCES

[1] P. Agashe, R. Rezaiifar, and O. Bender, "CDMA2000 High Rate Broadcast Packet Data Air Interface Design," *IEEE Communications Magzine*, 2004.

[2] P. K. Gopala and H. E. Gamal, "On the Throughput-delay Tradeoff in Cellular Multicast," in *Proc. of International Conference on Wireless Networks, Communications and Mobile Computing*, June 2005.

[3] U. C. Kozat, "On the Throughput Capacity of Opportunistic Multicasting with Erasure Codes," in *Proc. of IEEE INFOCOM*, 2008.

[4] G. Zeng, B. Wang, Y. Ding, L. Xiao, and M. Mutka, "Multicast Algorithms for Multi-Channel Wireless Mesh Networks," in *IEEE International Conference on Network Protocols (ICNP)*, 2007.

[5] J. Jin and B. Li, "Cooperative Multicast Scheduling with Random Network Coding in WiMAX," in *Proc. of IEEE International Workshop on Quality of Service (IWQoS)*, 2009.

[6] S.-E. Elayoubi, B. Fourestie, and X. Auffret, "On The Capacity of OFDMA 802.16 Systems," in *Proc. of IEEE International Conference on Communications (ICC '06.)*, June 2006.

[7] *MIMO Wireless Communications: From Real-World Propagation to Space-Time Code Design*. Academic Press, 2007.

[8] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "802.11 with Multiple Antennas for Dummies," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 1, pp. 19 – 25, 2010.

[9] M. Alicherry, R. Bhatia, and L. Li, "Joint Channel Assignment and Routing for Throughput Optimization in Multi-Radio Wireless Mesh Networks," in *Proc. of ACM Mobicom*, 2005.

[10] M. Kodialam and T. Nandagopal, "Characterizing the Capacity Region in Multi-Radio Multi-Channel Wireless Mesh Networks," in *Proc. of ACM Mobicom*, 2005.

[11] T. C.-Y. Ng and W. Yu, "Joint Optimization of Relay Strategies and Resource Allocations in Cooperative Cellular Networks," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, February 2007.

[12] H. Xu and B. Li, "XOR-Assisted Cooperative Diversity in OFDMA Wireless Networks: Optimization Framework and Approximation Algorithms," in *Proc. of IEEE INFOCOM*, 2009.

[13] S. Deb, V. Mhatre, and V. Ramaiyan, "WiMAX Relay Networks: Opportunistic Scheduling to Exploit Multiuser Diversity and Frequency Selectivity," in *Proc. of ACM Mobicom*, 2008.

[14] A. Gopinathan, Z. Li, and C. Williamson, "Optimal Multicast in Multi-Channel Multi-Radio Wireless Networks," in *Proc. of IEEE MASCOTS*, 2009.

[15] E. Gelal, K. Pelechrinis, I. Broustis, S. V. Krishnamurhty, S. Mohammed, A. Chockalingam, and S. Kasera, "On the Impact of MIMO Diversity on Higher Layer Performance," in *Proc. of IEEE ICDCS*, 2010, pp. 764–773.

[16] G. Song and Y. G. Li, "Cross-layer Optimization for OFDM Wireless Network - Part I and Part II," *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, March 2005.

[17] D. Wu, Y. T. Hou, and Y. Zhang, "Scalable Video Coding and Transport over Broad-Band Wireless Networks," *Proceedings of the IEEE*, 2001.

[18] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, "Trading Structure for Randomness in Wireless Opportunistic Routing," in *Proc. of ACM SIGCOMM*, 2007.

[19] T. Ho, M. Medard, J. Shi, M. Effros, and D. Karger, "On Randomized Network Coding," in *Allerton Conference on Communication, Control, and Computing*, October 2003.

[20] H. D. Sherali and L. Liberti, "Reformulation-Linearization Methods for Global Optimization," *Journal of Global Optimization*, 1991.

[21] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[22] "IEEE Std. 802.16e-2005. IEEE Standard for Local and metropolitan area networks, Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems," December 2005.