

Network Modulation: A New Dimension to Enhance Wireless Network Performance

Zhe Yang, Yuanqian Luo, and Lin Cai
Dept. of Electrical & Computer Engineering
University of Victoria, Victoria, BC Canada

Abstract—We introduce an approach called *network modulation* which gives us a new dimension to improve wireless network throughput and save energy. In current wireless systems, when a source transmits data to the receiver through a single-hop or multi-hop wireless path, the physical layer modulates and demodulates the information bits hop-by-hop, and the transmission over each hop is treated the same as in a point-to-point communication link. Given the broadcast nature of wireless medium and the wide variation of wireless channel quality, we let a sender transmit messages to multiple receivers simultaneously, using a software mapping technology, called network modulation, to redefine the constellation of typical quadrature amplitude modulation (QAM) schemes. As the software-based network modulation schemes do not require specialized communication hardware, they can be implemented with low cost and high flexibility. Network modulation can be used to improve network performance in a wide range of scenarios, for anycast (broadcast, multicast and unicast) services, one-way or two-way traffic, and single-hop or multi-hop wireless paths, in infrastructure or ad hoc networks. The minimum requirement for applying network modulation is that there are no less than three nodes within each others' transmission ranges, so we can consider modulation, topology control, resource allocation, and routing jointly.

I. INTRODUCTION

With the ever-increasing demand for Internet access anywhere, any time, and the limited wireless spectrum and energy supply, how to improve the throughput of wireless networks with the energy constraint is a key, pressing issue. Different from the point-to-point wired links, wireless medium is shared and the communication is inherently multi-point to multi-point. Therefore, wireless communication technologies and networking solutions are naturally coupled.

Inspired by this, extensive cross-layer designs for wireless networks have been proposed, e.g., how to use the physical (PHY) layer information to optimize (sometimes jointly) link layer resource allocation, network layer routing and transport layer end-to-end congestion control [1], [2]. Another direction is to extend the functions of the PHY layer to networking protocols, which can effectively enhance system performance. For example, error coding has been investigated for the PHY layer for several decades. Recently, network coding has been proposed and been applied to improve wired and wireless network performance [3]–[6].

Different from the previous approaches, in this paper, we let the PHY layer be adaptive to the network topology and propose the *network modulation* (NM) that can significantly improve network throughput with a software upgrade to wireless devices. In other words, we can use the existing

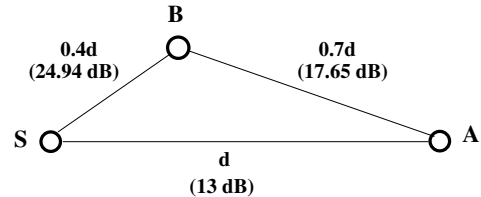


Fig. 1. A three-node example.

transceiver hardware to achieve the performance gain. The name network modulation is given as we combine topology control, scheduling and routing (which belong to the functions of network protocols) with modulation for multi-point wireless systems.

We use a three-node topology shown in Fig. 1 to brief the idea. Nodes S, A and B communicate with each other using wireless transmissions, and the transceiver distances and the corresponding signal to noise ratios (SNRs) are given in the figure, assuming the transmission power (symbol energy) is a constant, E_s . In current broadband wireless communication systems, the adaptive modulation and coding (AMC) schemes are used to improve transmission efficiency in highly dynamic environment, so the sender can adjust the modulation and coding schemes (and thus the link data rate) according to the received SNR. For example, in WiMAX systems, the transmitter can vary the data rate by using BPSK, QPSK, 16-QAM, and 64-QAM, combined with different error control codes.

For this three-node topology, we consider three typical cases: (1) S unicasts data to A; (2) S broadcasts data to both A and B; (3) S unicasts different data to A and B, respectively. For the first case, if using the direct transmission link with AMC, S can transmit at most two bits per symbol duration (b/sym) (using QPSK); if using B as the relay, the first hop can transmit 6 b/sym (using 64-QAM) and the second hop can transmit 4 b/sym (using 16-QAM), so the total throughput from A to S is $1/(1/6+1/4) = 2.4$ b/sym. Using the proposed NM scheme, we can let S send two bits to A and three other bits to B simultaneously during one symbol transmission (in Section IV, we will discuss how NM can achieve it), then let B relay the three bits to A using 16-QAM, so it takes $1+3/4$ symbol durations to transmit five bits to A and the throughput is $5/(1+3/4) = 2.857$ b/sym. The per-bit energy consumption using direct transmission, relay, and NM are $E_s/2$, $E_s/2.4$, and $E_s/2.857$, respectively. In other words, the proposed NM

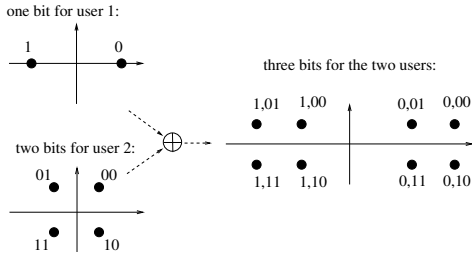


Fig. 2. Super-positioning precoding (SPC).

scheme can achieve both 43% and 19% throughput gains and 30% and 16% energy savings than the direct transmission and the traditional relay transmission schemes, respectively. The throughput gain for case (2), the broadcast scenario, is the same as that for case (1).

For case (3), assume that the lengths of the two messages to A and B are the same. If S transmits to A and B directly, it takes $1/2 + 1/6 = 2/3$ symbol durations to transmit one bit to A and one bit to B separately, i.e., the system throughput is 3 b/sym. If S transmits all bits to B first using 64-QAM, and let B relay the bits for A using 16-QAM, then the throughput is $2/(1/6 + 1/2.4) = 3.429$ b/sym. Using the proposed NM scheme, S can transmit two bits to A and three bits to B simultaneously in one symbol; then, for the remaining bits to A, we can use B as the relay using the same NM scheme as that for case (1). The throughput using NM is $6/(1 + 1/2.857) = 4.444$ b/sym. Thus, the throughput gains of using the proposed NM over the direct transmission and traditional relay schemes are 48% and 30%, respectively; and the energy savings are 32% and 23%, respectively.

The main contributions of this paper are threefold. First, we propose and promote the network modulation to improve wireless network throughput and save energy. It provides one more dimension of freedom for network optimization. The NM can be applied to many network scenarios, including unicast and multicast transmissions in infrastructure-based and ad hoc networks. Second, we investigate the performance of NM and propose how to optimize its configuration in different scenarios. Third, we further reveal the impact of topology control for applying the NM schemes in a network, i.e., how to choose relay nodes and how to group users to enhance the overall network performance.

The rest of the paper is organized as follows. In Section II, we introduce the background and the related work. The system model is presented in Section III. In Section IV, we propose the network modulation schemes. How to apply NM in different scenarios and optimize its configuration are studied in Section V, followed by performance evaluation by simulation in Section VI. Concluding remarks and further research issues are given in Section VII.

II. RELATED WORK

In 1970's, Thomas M. Cover proposed a super-positioning precoding (SPC) scheme to enhance the spectrum efficiency of broadcast channels [7]. The idea can be illustrated using

an example shown in Fig. 2. When a source needs to deliver information bits to two users with different channel quality, instead of transmitting the bits in different time slots (called the time-sharing scheme), the source can superposition the bits targeted to different users, and transmit them in one shot, such that both users can demodulate and decode its own bits successfully (w.h.p.). From the information theoretical perspective, SPC can achieve higher total rates than the time-sharing scheme.

The potential of SPC has been realized not until recently, when the digital video broadcasting (DVB) systems have standardized and implemented the hierarchical modulation (h-mod) schemes based on the idea of SPC [8], [9]. The nice fit of SPC (or h-mod) and video broadcasting is because the video sources use scalable video coding schemes and prefer differentiated services (DiffServ) provided by SPC. For instance, each video frame is encoded to a base layer and a number of enhancement layers, and bits of the base layer are much more important. Using SPC (or h-mod), the base layer bits can be received successfully with much lower SNR than that for the enhancement layer bits, so the receiver can always decode the base layer successfully, and decode the enhancement layers when the transmission channel quality is sufficiently good. However, SPC (or h-mod) requires specialized hardware and sophisticated signal processing, so it is not desirable for wireless devices.

In [10], we designed a scalable modulation (s-mod) scheme which uses software to remap the constellations of typical QAM schemes to provide DiffServ for video multicast over wireless networks. Performance study shows that the software based solution, s-mod, can achieve performance similar to or even better than SPC (or h-mod) for supporting video multicast in dynamic wireless environment. SoftCast is another approach to support scalable video multicast in wireless networks, by setting the transmitted signal proportional to the video pixel value [11].

In this paper, we use a software based modulation solution, similar to s-mod, to improve the wireless network performance. Different from the work in [10], [11], here, we focus on how the software based modulation remapping can be combined with the link layer resource allocation and network layer relay and routing to significantly improve the throughput of wireless networks, for any types of applications (with or without DiffServ requirements), anycast services, and single-hop or multi-hop wireless networks.

There are many other approaches in different layers to improve wireless network performance. User cooperative diversity (CD) systems can emulate the multiple transmit antennas situation in multiple input and multiple output (MIMO) systems to achieve spatial diversity gains [12]–[19]. As cooperative diversity systems require complicated bit-level coordination and signal processing, we do not incorporate CD in this work. How to combine CD and NM to further enhance network performance will be an interesting further research issue.

Packet-level, symbol-level, and bit-level network coding

(NC) can utilize the broadcast nature of the wireless channel to improve the system throughput with certain topologies and traffic patterns [3]–[5]. The NM schemes proposed in this paper can be beneficial in many general cases. NM and NC are orthogonal approaches, and it is possible to utilize both of them simultaneously.

III. SYSTEM MODEL

For a wireless communication link, the raw data rate relies on the symbol rate (baud rate), modulation and coding schemes. For simplicity, we assume that all links use the same symbol rate and error code. Hence, the link data rate is determined by the modulation scheme employed. AMC has been widely deployed in broadband wireless communication systems. According to the channel quality (received SNR), the modulation type is chosen to maximize the data rate under the constraint that the bit error rate (BER) is below a threshold. In this paper, we also adopt the greedy adaptive modulation selection scheme, so the sender always chooses the highest applicable modulation scheme under the BER constraint.

Wireless signal suffers from path-loss, shadowing, multipath fading and other impairments. The transceiver distance d is assumed fixed during the period of each transmission, so the path-loss can be estimated by

$$PL(d)[\text{dB}] = PL(d_0)[\text{dB}] + 10 \cdot \alpha \log_{10}\left(\frac{d}{d_0}\right), \quad (1)$$

where α is the path-loss exponent, and $PL(d_0)$ is the path-loss at reference distance d_0 . $PL(d_0)$ can be calculated by the Friis free-space model:

$$PL(d_0)[\text{dB}] = 10 \cdot \log_{10}\left(\frac{G_{tx}G_{rx}v^2}{(4\pi)^2 d_0^2 L}\right), \quad (2)$$

where v is the wavelength corresponding to the center frequency f_c , L is the system loss factor, and G_{tx} and G_{rx} are the transmitter and receiver antenna gains, respectively.

The received signal power P_r can be expressed as

$$P_r = \kappa P_t \left(\frac{d}{d_0}\right)^{-\alpha}, \quad (3)$$

where $\kappa = 10^{PL(d_0)/10} \beta$ equals the reference path-loss times the fading and shadowing factor β , and P_t is the transmission power. The proposed NM can be applied when different links have different values of β . For simplicity, we assume $\beta = 1$ for all links in the following calculation. Then, the received SNR, γ , is given by

$$\gamma = \frac{P_r}{N_0} = \frac{\kappa P_t d^{-\alpha}}{N_0} = \gamma_0 \left(\frac{d}{d_0}\right)^{-\alpha}, \quad (4)$$

where γ_0 is the SNR at the reference distance d_0 .

IV. NETWORK MODULATION

A. Motivation

The design of network modulation is inspired by the example shown in Fig. 1. In this example, α is three. Given the BER requirement of 10^{-3} , the received SNR thresholds to support QPSK, 16-QAM, and 64-QAM are 9.8 dB, 16.8 dB,

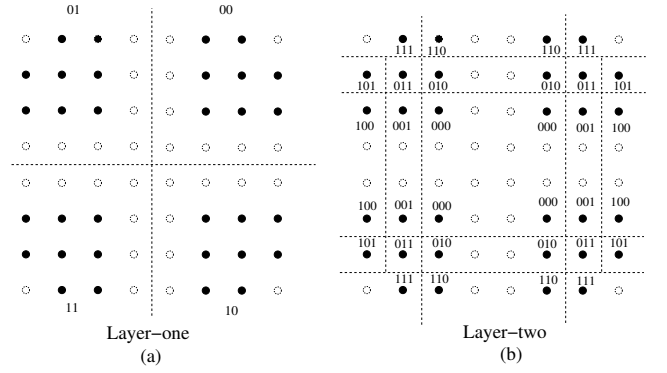


Fig. 3. A network modulation design.

and 22.7 dB, respectively. The received SNR of SA (link between node S and node A) is 13 dB which can support QPSK (2 b/sym), the received SNR of SB is 24.94 dB which can support 64-QAM (6 b/sym), and the received SNR of BA is 17.65 dB which can support 16-QAM (4 b/sym). The first observation is that, to deliver information bits from node S to node A, the path loss between node S and node A is so severe that it is preferable to use node B as a relay. This relay strategy has been widely used in relay networks and the multi-hop routing protocol design.

The second observation is that, if using node B as the relay, when node S transmits bits to node B using 64-QAM, node A can receive a copy of the information. But as the channel quality of SA is so poor, node A may not be able to demodulate any bit successfully from this transmission. In the PHY layer, user cooperative communication has been extensively studied, which tries to use sophisticated signal processing techniques to combine the two copies of the signal node A received (the one from the transmission by node S and the other from the transmission by node B) to enhance system performance. Here, we found that, even without such techniques which require hardware upgrade, we can enhance the relay performance by re-map the constellation of the 64-QAM used by node S, as shown in Fig. 3.

Using the re-defined constellation map in Fig. 3, each symbol contains five bits information represented by the location of the 32 black points. These 32 points belong to four clusters, and the index of the clusters represent two layer-one bits, as shown in Fig. 3 (a). Each cluster contains eight points, which represent three layer-two bits, as shown in Fig. 3 (b). The dotted lines in Fig. 3 (a) and (b) give the demodulation decision boundaries of the layer-one and layer-two bits, respectively. As the minimum distances from the constellation points to the layer-one decision boundaries are three times that to the layer-two decision boundaries, the minimum SNR needed to transmit the layer-one bits successfully (which is close to the SNR needed for QPSK) is much less than that for the layer-two bits (close to the SNR needed for 64-QAM).

Therefore, when node S transmits a symbol using the NM design shown in Fig. 3, node A can demodulate two layer-one bits successfully, and node B can demodulate both the

two layer-one bits and three layer-two bits successfully. Then, node B just needs to relay the layer-two bits to node A using 16-QAM (with 4 b/sym). Node A can receive five bits with $1 + 3/4$ symbol duration, so the throughput using NM-assisted relay is $5/1.75 = 2.857$ b/sym, which is higher than that with both direct transmission (2 b/sym) and purely relying on node B to relay all bits (2.4 b/sym).

As we use the same energy to transmit each symbol, the energy consumed per bit is inversely proportional to the throughput. Thus, in the above example, using NM-assisted relay can enjoy both throughput gain and energy saving.

B. NM Implementation

To implement NM schemes, assume node S and node B having the knowledge of channel quality between SB, SA and BA. In the sender side, we can add a mapping function before modulation. Using the above example, node S groups five bits and maps them to a six-bit symbol according to the standard constellation map of 64-QAM. For example, “00000” is mapped to “000000” in Gray-mapped 64-QAM constellation, and “00111” is mapped to “001101”.

The software re-mapping is so flexible that we can arrange the bit-mapping in different clusters in an arbitrary order such that the decision regions of each symbol can be maximized. For example, we can put the layer-two bits “000” in the neighboring clusters closest to each other as shown in Fig. 3 (b). This strategy can slightly enhance the BER performance of layer-two bits.

In the receiver side, both node A and node B can demodulate the symbols using standard 64-QAM demodulator first. Then, node A re-maps each 6-bit symbol to two (layer-one) bits, and B re-maps each 6-bit symbol to three (layer-two) bits.

Therefore, the implementation of NM (for both modulation and demodulation) can rely on software upgrade using the existing standard QAM modems.

C. BER Estimation

The BERs for layer-one and layer-two bits are determined by the probabilities that the received signal is outside their decision regions, respectively.

Using an AWGN channel as an example, we can integrate the 2-D Gaussian distribution over the area of a layer-one symbol’s decision region to get the probability that the corresponding layer-one bits are successfully demodulated, p_s . If this symbol contains m layer-one bits, the BER of layer-one bits is approximately $(1 - p_s)/m$. Similarly, we can obtain the BER of layer-two bits.

Thus, when designing NM to ensure the BER performance, a key issue is to set the minimum Euclidean distance (MED) of the constellation points to the decision boundaries appropriately. On the other hand, the average symbol energy is proportional to the square of the Euclidean distance of the constellation points to the origin. Therefore, when we do the re-mapping, we should avoid the points farthest away from the origin if possible. The design given in Fig. 3 has used these two insights to achieve performance gains.

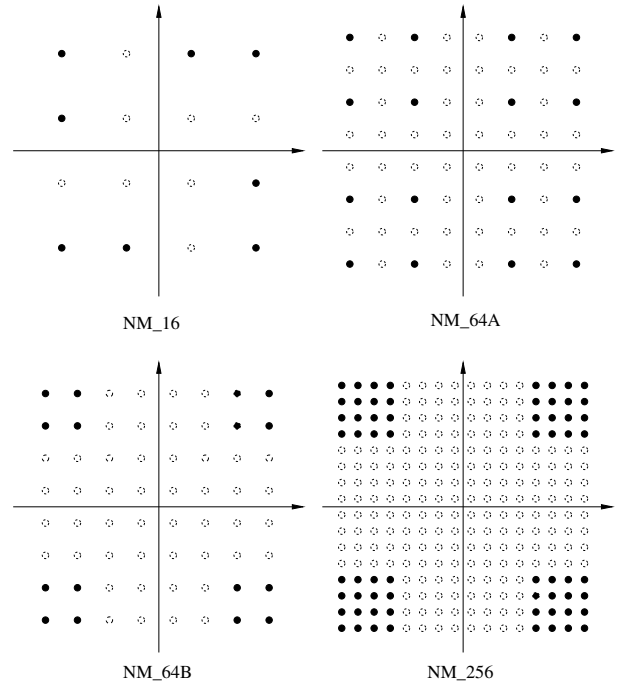


Fig. 4. Network Modulation Schemes.

D. NM Schemes

In this paper, in addition to the NM scheme shown in Fig. 3, we design four other NM schemes shown in Fig. 4. We use these five NM schemes to study how effectively we can improve the network performance. These NM schemes are based on 16-QAM, 64-QAM, and 256-QAM, and each symbol contains two layer-one bits (with larger MED) and one to four layer-two bits (with smaller MED).

Note that we have NOT exhaustively searched the design space of NM yet, which remains an open issue. All existing standard modulations can be viewed as special cases of NM.

V. NM APPLICATIONS

In this section, we study how to apply NM in various network scenarios to achieve performance gains.

A. Relay

The first scenario is that a source node delivers information to a destination, with the help of another node which serves as a relay. In this subsection, we first give the necessary conditions that using a relay is preferable, with or without NM. The conditions can be used to identify relay candidates. Then, we propose how to optimally configure NM.

1) *Necessary condition of using a relay:* With traditional modulation, the direct transmission between the source and the destination can support up-to m b/sym. We have the following theorem.

Theorem 1: It is preferable to use a relay with traditional modulation if the following condition holds:

$$\frac{m_1 m_2}{m_1 + m_2} > m, \quad (5)$$

where m_1 and m_2 are the data rates (b/sym) corresponding to the highest modulation schemes that can be supported by the source to the relay and the relay to the destination, respectively.

For a NM scheme, the source can transmit m_1' bits and m_2' bits to the destination and the relay in one symbol, respectively. The following theorem gives the necessary condition that the NM-assisted relay is preferable.

Theorem 2: The condition that using the NM-assisted relay is preferable to the direct transmission is

$$\frac{(m_1' + m_2')m_2}{m_2 + m_2'} > m. \quad (6)$$

The LHSs of (5) and (6) give the throughputs (in terms of bits per symbol) of the relay with traditional modulation and with NM, respectively. The proofs of the above two theorems are straight-forward and are omitted here.

2) *Optimal NM configuration:* Given an identified relay candidate, the next question is how to choose the best NM scheme under the constraint of BER, p_e . Denote $p_{e,sr}$, $p_{e,rd}$, and $p_{e,sd}$ the BERs of the source (S) to the relay (R), R to the destination (D), and S to D, respectively, which depend on the modulation schemes used and the transmission power. Assume that all nodes transmit using the same power P_t . We can formulate the following optimal NM configuration problem.

Problem 1: (P1)

$$\begin{aligned} \max \quad & \frac{(m_1' + m_2')m_2}{m_2 + m_2'}, \\ \text{s.t.} \quad & p_{e,sr}(m_2', P_t) + p_{e,rd}(m_2, P_t) \leq p_e \\ & p_{e,sd}(m_1', P_t) \leq p_e \end{aligned}$$

As the number of modulation schemes that we can choose is limited (in this paper, we consider the five proposed NM schemes and four traditional QAM modulation schemes), it is feasible to solve the above optimization problem by exhaustive searching using Algorithm 1.

Algorithm 1 NM_Configuration_Relay

Require: (p_e , P_s , P_r) { p_e is the required BER. P_s and P_r are the transmission powers of the source node and the relay node, respectively.}

- 1: estimate the received SNRs for SD, SR, and RD channels
 - 2: $R = 0$
 - 3: **for** all NM schemes **do**
 - 4: calculate $p_{e,rd}$, $p_{e,sr}$ and $p_{e,sd}$;
 - 5: **if** $p_{e,sr} + p_{e,rd} \leq p_e$ and $p_{e,sd} \leq p_e$ **then**
 - 6: **if** $\frac{(m_1' + m_2')m_2}{m_2 + m_2'} > R$ **then**
 - 7: $R = \frac{(m_1' + m_2')m_2}{m_2 + m_2'}$
 - 8: set n equal to the index of the NM scheme
 - 9: **end if**
 - 10: **end if**
 - 11: **end for**
 - 12: **Return:** R , n
-

3) *Optimal power allocation:* As the channels of SR and RD might be quite different, appropriate transmission power allocation might be desirable to further enhance the system performance. Denote the total power budget by P . The optimal

configuration considering power allocation can be formulated as the following optimization problem.

Problem 2: (P2)

$$\begin{aligned} \max \quad & \frac{(m_1' + m_2')m_2}{m_2 + m_2'} \\ \text{s.t.} \quad & p_{e,sd}(m_1', P_1) \leq p_e \\ & p_{e,sr}(m_2', P_1) + p_{e,rd}(m_2, P_2) \leq p_e \\ & P_1 + P_2 \leq P_{\max} \\ & P_1 \geq 0, \quad P_2 \geq 0 \end{aligned}$$

The optimization problem (P2) is a mixed-integer problem which is difficult to solve, especially without the close-form expression of BER. Considering that the power control in a practical system is discretized, P_1 and P_2 are discrete variables with limited number of values. Using exhaustive searching is still tractable as the searching space is moderate. We set the number of power levels to be 100 in this paper, which is sufficient for a practical system. We then use Algorithm 2 to solve the optimal power allocation problem (P2) based on Algorithm 1.

Algorithm 2 NM_Configuration_Power_Allocation

Require: (p_e , P) { p_e is the required BER and P is the total power budget.}

- 1: $R = 0$;
 - 2: **for** $P_1' = 0$ to P **do**
 - 3: $P_2' = P - P_1'$
 - 4: $(R', n) = \text{NM_Configuration_Relay}(p_e, P_1', P_2')$
 - 5: **if** $R' > R$ **then**
 - 6: $R = R'$
 - 7: $P_1 = P_1'$; $P_2 = P_2'$
 - 8: **end if**
 - 9: **end for**
 - 10: **Return:** P_1 , P_2 , R , n
-

B. Broadcast and Multicast

When a node broadcasts or multicasts to a group of receivers using wireless transmissions, it needs to adjust the modulation so the receiver with the worst channel quality can successfully receive the message. In this case, we can use the similar strategy as discussed in the previous subsection to enhance the system performance. That is, first identify relay candidates for the worst-channel receiver; then use the optimal NM configuration algorithm to select the appropriate NM scheme to maximize the throughput and save energy. Then we repeat the procedure to consider the second-worst channel receiver, and so on. Due to the space limitation, we do not further discuss this problem here.

C. Downlink Unicast

Next, we consider the scenario that, in an infrastructure-based network, a base station (BS) or an access point (AP) needs to send different messages to users using the downlink. We can also use NM to save the downlink bandwidth and energy. If we let an idle node act as a relay for an active user, the problem is again the same as the relay problem discussed

in subsection V-A, and we omit it here. If only active nodes can collaborate with each other, the problem is different, which is the focus of this subsection.

Let \mathcal{F} be the active user set (the downlink receiver set), and each user needs to receive one unit of data per scheduling period from the BS. Using NM, the BS can transmit messages to multiple users with different channel quality simultaneously to achieve performance gain. As a starting point, we use NM to combine the messages to two users only, and how to combine the messages sending to more than two users remains a further research issue.

The propose strategy takes two steps. First, given a pair of users, select the optimal NM scheme for them. Second, group active users into pairs appropriately to maximize the whole network throughput.

1) *NM configuration for a collaborative pair of users:*

For a pair of users i and j , without loss of generality, we assume that the channel from the BS to user i is better than that from the BS to user j . Let m_i and $m_{i,j}$ be the highest data rates (b/sym) using the traditional modulation types for the channels from the BS to user i and from user i to user j , respectively. For the transmission from the BS to user i and user j using NM, we denote m'_i and m'_j the data rates (b/sym) for user i and user j , respectively. We can obtain the optimal NM configurations using Algorithm 3.

Algorithm 3 NM_Configuration_Downlink_Cooperation

Require: (p_e, P_s, P_i) $\{p_e$ is the required BER. P_s is the transmission power of the base station. P_i is the transmission power of user i to relay for user $j\}$

```

1:  $R = 0$ 
2: for all NM schemes do
3:   calculate  $p_{e,si}$  and  $p_{e,sj}$ 
4:   if  $p_{e,si} \leq p_e$  and  $p_{e,sj} \leq p_e$  then
5:     if  $m'_i \leq m'_j$  then
6:       if  $\frac{m'_j}{1+(m'_j-m'_i)/m_i} > R$  then
7:          $R = \frac{m'_j}{1+(m'_j-m'_i)/m_i}$ 
8:         set  $n$  to the index of the NM scheme
9:         set  $n'$  to the index of the  $m_i$  modulation scheme
10:      end if
11:    else
12:       $(R_0, n_0) = \text{NM\_Configuration\_Relay}(p_e, P_s, P_i)$   $\{\text{find the best NM scheme to relay message to } j \text{ by } i\}$ 
13:
14:      if  $\frac{m'_i}{1+(m'_i-m'_j)/R_0} > R$  then
15:         $R = \frac{m'_i}{1+(m'_i-m'_j)/R_0}$ 
16:        set  $n$  to the index of the NM scheme
17:        set  $n' = n_0$ 
18:      end if
19:    end if
20:  end for
21: end for
22: Return:  $R, n, n'$ 

```

For a NM scheme which can transmit m'_i bits to user i and m'_j bits to user j per symbol, if $m'_i \leq m'_j$, the BS can transmit the remaining bits to user i directly using a traditional

modulation (lines 5-10 in Algorithm 3); if $m'_i > m'_j$, the remaining bits to user j can be transmitted using i as the relay, and we can use Algorithm 1 to find the appropriate NM scheme (with index n') for the relayed bits (lines 11-18). Therefore, Algorithm 3 returns the total throughput using the best NM schemes (R), the NM scheme selected for the unicast transmission (n), and that for the relay transmission (n'). If we need to consider power allocation, we can simply add another loop to find the best power allocation strategy and the NM schemes.

2) *Matching algorithm:* For the downlink transmission, one key issue of applying NM is how to group users to maximize the total downlink throughput. This can be formulated as a non-bipartite weighted matching problem.

Denote by $R_{i,j}$ the throughput for a pair of users i and j with the optimal NM configurations, and R_i and R_j their throughputs with traditional modulations, respectively. The gain of using NM for users i and j is $G_{i,j} = \max(1/R_i + 1/R_j - 2/R_{i,j}, 0)$. (If $2/R_{i,j} \geq 1/R_i + 1/R_j$, we do not use NM for the pair, and the gain is zero in this case.) The optimal matching will maximize the summation of $G_{i,j}$ for all users.

If $G_{i,j}$'s for all users in \mathcal{F} are known, the state-of-the-art optimal matching algorithm can solve the above matching problem with the complexity of $O(n^3)$ [20]. However, using Algorithm 2 to obtain $G_{i,j}$'s for all $(n^2/2)$ pairs is non-trivial.

Therefore, instead of using the optimal matching algorithm, we modify a simple heuristic matching algorithm called worst-link-first (WLF) matching [21] to group the users:

- (1) The BS first sorts the channel qualities of all users.
- (2) The BS selects an unmatched user j with the worst channel quality among all of the unmatched users.
- (3) The BS selects an unmatched user i such that $G_{i,j}$ (obtained using Algorithm 3) is the maximum one among all $G_{k,j}$, where k is an unmatched user other than j . If $G_{i,j} > 0$, i and j are added to the matched set; otherwise, j is added to the matched set.
- (4) Repeat (2) and (3) until the number of unmatched users is less than two.

The WLF matching algorithm above has the computational complexity of $O(n^2)$ only. More importantly, it does not require to pre-calculate the NM gains of all possible pairs. The most time-consuming step (3) of the WLF algorithm can be further simplified by using the knowledge of the network topology, i.e., the BS can choose i from those located in the desirable relay region of user j , which will be discussed in Section VI.

D. Uplink Unicast

The uplink scenario is similar to the downlink case. A user may need to rely on other nodes (who can be either active or inactive) to relay messages for it. We can also use the WLF matching algorithm to group nodes into pairs if using NM is preferable. The NM configuration problem for each pair can also be solved using Algorithms 1 or 2, depending on whether power allocation is used or not.

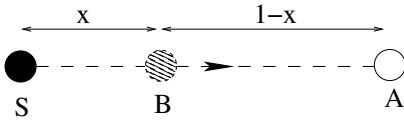


Fig. 5. Line topology.

VI. PERFORMANCE EVALUATION AND DISCUSSION

Extensive simulations have been conducted to evaluate the performance of NM in different scenarios with different network topologies. For comparison purpose, we also implement the traditional relay scheme based on adaptive modulation. We use the same average symbol energy for all transmission schemes. Thus, if a scheme can achieve a higher throughput, it can also save energy (using less energy per bit). In this section, we focus on the throughput gain only, which can be easily converted to energy saving.

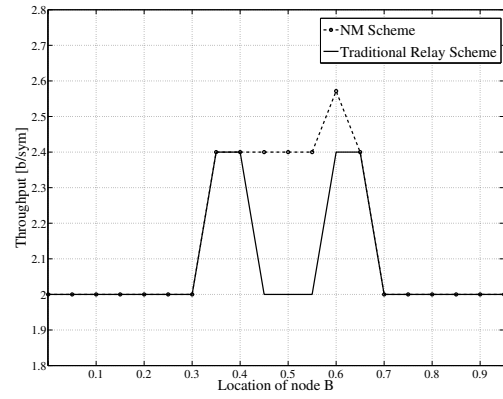
A. Line Topology

We first consider a line topology, where three nodes S, A, and B, are located on a line segment, as shown in Fig. 5. An example for this topology is in vehicular networks where vehicles (nodes) are located on streets (lines).

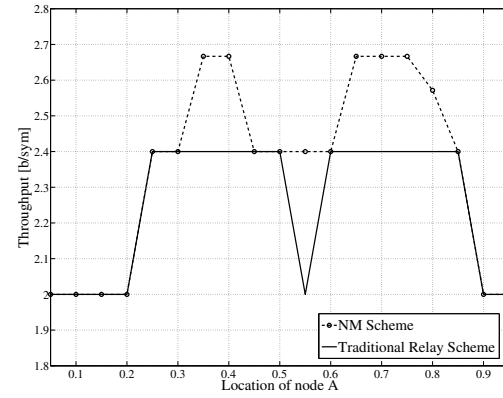
1) *Simulation setting*: The distance of node S and node A is fixed at $d_0 = 1$ unit with the corresponding reference SNR equal to 15 (11.76 dB). α is set to three, so the received SNR at distance d equals $15(d/d_0)^{-3}$. The BER requirement is $p_e = 10^{-3}$, and the AWGN channel model is used in the simulation. Using direct transmission, node S (located at $x = 0$) can transmit to node A (located at $x = 1$) at most 2 b/sym using QPSK. The location of node B is between node S and node A.

2) *Simulation results*: First, we let node S transmit to node A, using node B as a relay. In Fig. 6 (a) and (b), we compare the throughputs without and with the optimal power allocation. In this figure, x-axis represents the location of node B, and y-axis is the throughput (i.e., the number of bits per symbol received by node A). As shown in Fig. 6 (a) and (b), first, to exceed the throughput of direct transmission (2 b/sym), the location of B can be in a wider range using the NM-assisted relay schemes than that using the traditional relay. Second, the proposed NM with a fixed transmission power can outperform the traditional relay schemes (by up-to 20%) for $x \in (0.4, 0.7)$; if using optimal power allocation for all schemes, the NM schemes can outperform the traditional relay schemes with an even wider range of x . The results demonstrate the importance of power allocation. In the following, all the results are obtained with the optimal power allocations. If the SNR of SA changes (e.g. the case shown in Fig. 1), we can achieve even higher performance gain.

Next, let node S transmit different messages to both node A and node B. The throughput results are given in Fig. 7. In this scenario, the NM can achieve even higher gains than the traditional schemes. From the figure, the NM schemes can achieve up-to 64% and 33% throughput gains than the traditional schemes without or with relay, respectively. In the



(a) Fixed power allocation



(b) Optimal power allocation

Fig. 6. Throughput comparison, single-receiver.

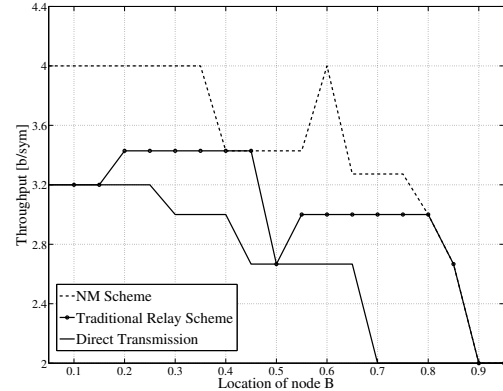


Fig. 7. Throughput comparison, two-receiver.

above figures, the throughput curves are in zig-zag patterns. This is because the number of modulation schemes is limited. A change of location of node B may change the rates of both the first hop (SB) and the second hop (BA). For instance, with equal power allocation, when $x = 0.5$, both hops can support 16-QAM only, so the throughput with traditional relay is 2 b/sym; when $x = 0.35$ or 0.65 , one of the hop can support 64-QAM and the other hop can support 16-QAM, so the throughputs at both locations are 2.4 b/sym, higher than that at $x = 0.5$. This is different from the intuition

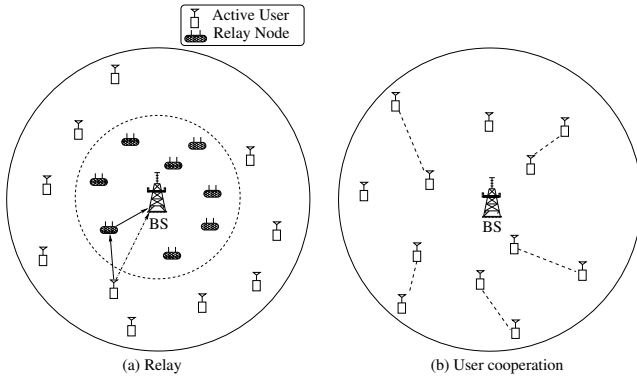


Fig. 8. Network topology.

that the best relay should always be located in the midpoint of the transceiver. In addition, when the received SNR of SA changes, the desirable relay regions also change, due to the limited choices of modulation schemes. Another observation is that the preferable relay regions (in which relay schemes can outperform direct transmission schemes) are different for the single-receiver case and the two-receiver case. This is because that the throughput of the single-receiver case is highly dependent on the rate of the second hop, which should support more than 2 b/sym in order to outperform the direct transmission schemes. Thus, when B is too close to S, using relay with or without NM cannot achieve higher throughput. For the two-receiver case, the transmissions are finished in two stages: unicast transmission and relay transmission (if the transmission for node B finishes first) or direct transmission (if the transmission for node A finishes first). In the unicast stage, the data for different nodes can be transmitted simultaneously, so the overall throughput is high. Thus, when node B is close to node S, even though the second hop channel condition is not good enough to support modulation schemes higher than QPSK, the NM-assisted transmission scheme can still achieve throughput gain. Only when node B is very close to node A, we cannot obtain any gain by traditional or NM-assisted relay schemes.

The results with the line topology not only demonstrate the performance gains of the NM, but also provide the important insight of where the most desirable relay nodes are located, which is useful for the relay selection or matching procedures.

B. Downlink and Uplink Transmissions

We further evaluate the performance in an infrastructure-based network, where a BS is located in the middle to serve all users in the cell. We consider two different scenarios as shown in Fig. 8: (a) besides the active users, there are dedicated nodes who can relay messages for active users, and (b) there is no dedicated relay node, and the active users collaborate and relay for each other. For the first scenario, \mathcal{N} is the set of relay nodes which are located in the region near the BS and \mathcal{F} is the set of active users which are located in the region further away from the BS. The number of nodes in \mathcal{N} is the same as that in \mathcal{F} , i.e., $|\mathcal{N}| = |\mathcal{F}|$. For the second scenario, active users are

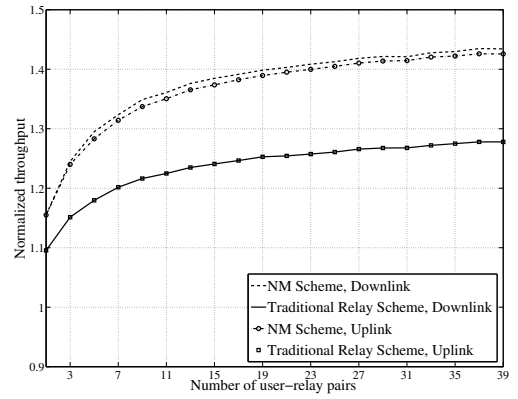


Fig. 9. Per-user throughput, with dedicated relay.

randomly distributed in the cell. For both scenarios, the BS has the knowledge of the network topology which is assumed to be stable during the scheduling period. The BS uses the WLF matching algorithm to group relay nodes and active users into pairs for the first scenario (with dedicated relay nodes), and group active users into pairs for the second scenario (without dedicated relay nodes). If using relay schemes for a pair cannot achieve any throughput gain, the direct transmission schemes will be used. We also use optimal power allocations for all schemes. For simplicity, we assume that there is no mutual interference for the downlink and uplink transmissions.

We tested different transmission schemes with different node density by varying the number of user-relay pairs (or user pairs) from 1 to 39 with the step size of 2. For each node density, we repeated the simulation 1000 times using different randomly generated topologies, and calculated the average results. For the scenario with dedicated relay nodes, the per-user throughputs for both uplink and downlink are shown in Fig. 9. Simulation results are normalized by the throughput of using direct transmission schemes.

In Fig. 9, we note that the downlink and uplink throughputs are the same with the traditional relay schemes. This is because, with the traditional relay schemes, the most preferable relay nodes are in the same or symmetric locations for the downlink and uplink transmissions. Thus, the same relay node will serve the relay for both uplink and downlink. However, the throughputs of downlink and uplink are not necessary to be the same with NM-assisted relay, since different relay nodes might be chosen for uplink and downlink transmissions. Thus, with NM, the work loads of relay nodes can be more balanced, which is desirable for energy constrained wireless networks. Overall, when the number of pairs is 39, the NM schemes can achieve 43% and 12% throughput gains than the direct transmission and the traditional relay schemes, respectively. For the second scenario, we consider the downlink only, as it is quite different from the relay cases studied above. The normalized average per-user throughputs are shown in Fig. 10. It is observed that when the number of user pairs is 39, the NM can achieve 43% and 30% throughput gains than the direct transmission and the traditional relay schemes, respectively.

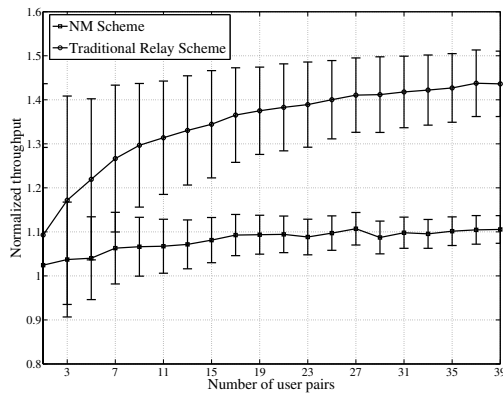


Fig. 10. Per-user throughput, user cooperation.

The error bars in the figure represent the standard deviations. From these figures, a common tendency for both scenarios is that the throughput increases slightly as the number of users increases. This is because, with higher density, it is easier to find a preferable relay or cooperative node for a user, so NM can have a better chance to be employed. Note that we only consider five NM schemes in this paper. If more well-designed NM schemes are deployed, it is possible to further improve the network performance.

VII. CONCLUSION AND FURTHER DISCUSSION

In this paper, we have proposed a new approach to improve wireless network performance by using *network modulation*. Based on software bit remapping, NM schemes can be implemented using existing transceiver hardware. Our results have shown that NM is applicable and desirable in many scenarios, with the minimum requirement that there are at least three nodes involved in the wireless communications. The idea behind NM is that we can optimize modulation considering the topology and setting of multiple nodes. Essentially, modulation, resource allocation, routing, and topology control can be optimized together.

The work reported in this paper opens up numerous further research opportunities. First, we need to re-define the admission region of single-hop and multi-hop wireless networks with NM, considering different types of traffic characteristics and QoS requirements. Second, the interactions of NM and the protocols of all layers need further investigation. For example, how to effectively retransmit and recover packet errors and losses in wireless networks with NM; how to devise effective distributed routing algorithms considering the impact of NM. Third, in this work, the receiver demodulates and decodes bits from one transmission. It is possible to combine various user cooperative diversity techniques with NM to improve the system performance by fully utilizing the spatial diversity gain. Fourth, we use a simple AWGN channel model to evaluate the NM performance, and the performance study using more complicated fading models is an open issue. Fifth, with NM, the topology control problem becomes more complicated. On the other hand, combining NM and topology control can be a more effective way to further enhance wireless network

performance. Finally, other new NM schemes can be designed and many new optimization problems can be formulated, as NM provides one more dimension of freedom to improve network performance.

REFERENCES

- [1] S. Shakkottai, T.S. Rappaport, and P.C. Karlsson. Cross-layer design for wireless networks. *IEEE Communications Magazine*, 41(10):74–80, 2003.
- [2] L. Chen, S.H. Low, M. Chiang, and J.C. Doyle. Cross-layer congestion control, routing and scheduling design in ad hoc wireless networks. In *IEEE Infocom'06*, volume 6, 2006.
- [3] R. Ahlswede, N. Cai, S.-Y.R. Li, and R.W. Yeung. Network information flow. *IEEE Transactions on Information Theory*, 46(4):1204–1216, Jul. 2000.
- [4] Sachin Katti, Dina Katabi, Hari Balakrishnan, and Muriel Medard. Symbol-Level Network Coding for Wireless Mesh Networks. In *ACM SIGCOMM'08*, pages 401–412, Seattle, WA, August 2008.
- [5] S. Zhang, S. Liew, and P. Lam. Hot topic: physical-layer network coding. In *ACM MobiCom'06*, pages 358–365, New York, NY, USA, 2006.
- [6] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft. XORs in the air: practical wireless network coding. *IEEE/ACM Transactions on Networking (TON)*, 16(3):497–510, June 2008.
- [7] T. Cover. Broadcast channels. *IEEE Trans. on Information Theory*, 18(1):2–14, Jan. 1972.
- [8] EN 300 421 v1.1.2. Digital Video Broadcasting (DVB), Framing structure, channel coding and modulation for 11/12 GHz satellite services. ETSI Standard, Aug. 1997.
- [9] S. Wang and B. K. Yi. Optimizing enhanced hierarchical modulations. In *IEEE Globecom'08*, Nov./Dec. 2008.
- [10] L. Cai, Y. Luo, S. Xiang, and J. Pan. Scalable modulation for scalable wireless videocast. In *IEEE Infocom'10, mini-symposium*, San Diego, CA, USA, Mar. 2010.
- [11] D. Katabi, H. Rahul, and S. Jakubczak. SoftCast: One video to serve all wireless receivers. Technical Report MIT-CSAIL-TR-2009-005, CSAIL, MIT, Feb. 2009.
- [12] W. Su, A.K. Sadek, and K.J. Ray Liu. Cooperative communication protocols in wireless networks: performance analysis and optimum power allocation. *Wireless Personal Communications*, 44(2):181–217, 2008.
- [13] Y. Luo and L. Cai. Throughput maximization for user cooperative wireless systems with adaptive modulation. In *IEEE ICC'10*, 2010.
- [14] EG Larsson and BR Vojcic. Cooperative transmit diversity based on superposition modulation. *IEEE Communications Letters*, 9(9):778–780, 2005.
- [15] S. Sharma, Y. Shi, Y.T. Hou, H.D. Sherali, and S. Kompella. Cooperative communications in multi-hop wireless networks: Joint flow routing and relay node assignment. In *IEEE Infocom'10*, 2010.
- [16] J.N. Laneman, D.N.C. Tse, and G.W. Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Transactions on Information Theory*, 50(12):3062–3080, 2004.
- [17] J. Yang, M. Ghosh, P.R.N. America, and B. Manor. A cooperative modulation scheme for wireless relay networks. In *IEEE VTC2007-Spring*, pages 1628–1632, 2007.
- [18] Z. Lin, E. Erkip, and M. Ghosh. Rate adaptation for cooperative systems. In *Proc. IEEE Globecom'06*, 2006.
- [19] V. Mahinthan, L. Cai, J. W. Mark, and X. Shen. Maximizing cooperative diversity energy gain for wireless networks. *IEEE Trans. on Wireless Communications*, 6(6):2530–9, June 2007.
- [20] H. N. Gabow. An efficient implementation of edmonds' algorithm for maximum matching on graphs. *J. ACM.*, 23(2):221–234, April 1976.
- [21] V. Mahinthan, L. Cai, J.W. Mark, and X. Shen. Partner selection based on optimal power allocation in cooperative-diversity systems. *IEEE Transactions on Vehicular Technology*, 57(1):511–520, 2008.