

# Network Coding-Aware Routing in Wireless Networks

Sudipta Sengupta, *Senior Member, IEEE*, Shravan Rayanchu, and Suman Banerjee, *Member, IEEE*

**Abstract**—A recent approach—COPE, presented by Katti *et al.* (*Proc. ACM SIGCOMM 2006*, pp. 243–254)—for improving the throughput of unicast traffic in wireless multihop networks exploits the broadcast nature of the wireless medium through opportunistic network coding. In this paper, we analyze throughput improvements obtained by COPE-type network coding in wireless networks from a theoretical perspective. We make two key contributions. First, we obtain a theoretical formulation for computing the throughput of network coding on any wireless network topology and any pattern of concurrent unicast traffic sessions. Second, we advocate that routing be made aware of network coding opportunities rather than, as in COPE, being oblivious to it. More importantly, our model considers the tradeoff between routing flows close to each other for utilizing coding opportunities and away from each other for avoiding wireless interference. Our theoretical formulation provides a method for computing source–destination routes and utilizing the best coding opportunities from available ones so as to maximize the throughput. We handle scheduling of broadcast transmissions subject to wireless transmit/receive diversity and link interference in our optimization framework. Using our formulations, we compare the performance of traditional unicast routing and network coding with coding-oblivious and coding-aware routing on a variety of mesh network topologies, including some derived from contemporary mesh network testbeds. Our evaluations show that a route selection strategy that is aware of network coding opportunities leads to higher end-to-end throughput when compared to coding-oblivious routing strategies.

**Index Terms**—COPE, network coding, network coding-aware routing, opportunistic network coding, wireless broadcast scheduling, wireless link scheduling, wireless networks.

## I. INTRODUCTION

NETWORK coding is gaining popularity as a mechanism to increase the utilization of both wired and wireless networks. We explain the basic idea of network coding using a very simple example consisting of three wireless nodes as shown in Fig. 1(a). In the figure, node 1 wants to send a single packet ( $A$ ) to node 3, while node 3 wants to send a single packet ( $B$ ) to node 1. Due to transmission range limitations, both these paths go via node 2. Using standard techniques of packet forwarding, four wireless transmissions would be needed to complete these end-to-end packet transfers. The following are a possible sequence of these transmissions: i) packet  $A$  is transmitted by 1,

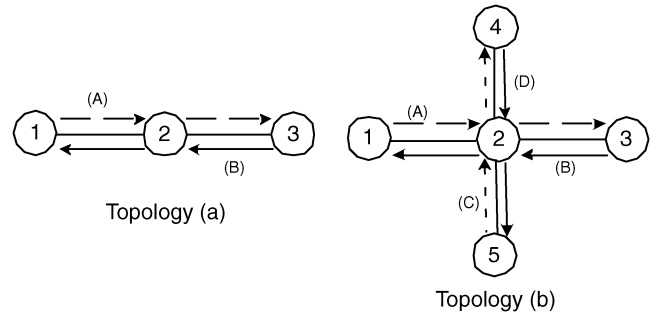


Fig. 1. Illustration of network coding.

with 2 being the intended recipient; ii) packet  $B$  is transmitted by 3, with 2 being the intended recipient; iii) packet  $A$  is transmitted by 2, with 3 being the intended recipient; and iv) packet  $B$  is transmitted by 2, with 1 being the intended recipient.

In comparison, using a simple form of network coding (as employed in the COPE approach [10]), the same two packets can be transferred by using three wireless transmissions instead of four using the following sequence: i) 1 transmits packet  $A$ , with 2 being the intended recipient; ii) 3 transmits packet  $B$  with 2 being the intended recipient; and iii) node 2 transmits a new packet  $C$  obtained by performing an XOR of packets  $A$  and  $B$ .

Both nodes 1 and 3 are intended recipients of this new packet. Wireless medium being inherently broadcast in nature allows such communication to be possible. Assuming node 1 still has a copy of packet  $A$ , it can obtain packet  $B$  by performing an XOR of packets  $A$  and  $C$ . Similarly, node 3 can obtain packet  $A$  by performing an XOR of packets  $B$  and  $C$ . Overall, this simple form of network coding achieved the same packet transfer effect on this two-hop path by using three transmissions instead of four and would lead to a throughput improvement of 33% in this example. Using prior terminology [10], we will refer to the original packets ( $A$  and  $B$ ) as *native packets* and the packet  $C$  derived as a combination of native packets as a *coded packet*.

In this paper, we focus on network coding as applicable to a multihop wireless network where there are *multiple concurrent unicast sessions*. We provide a theoretical framework for investigating the potential interactions between coding opportunities and routing decisions. Our goal is to develop techniques for systematically quantifying the benefits of network coding-aware routing across arbitrary wireless network topologies and traffic demands. We use a COPE-type network coding scheme for unicast traffic that exploits the broadcast nature of the wireless medium. More specifically, we study the following related set of issues in this paper.

**Estimating Coding Benefits:** Given *any* wireless topology, a set of traffic demands, and a coding-oblivious<sup>1</sup> routing strategy

<sup>1</sup>Coding-oblivious routing implies that routing decisions are not made based on coding opportunities available. This approach requires no change in existing routing mechanisms.

Manuscript received August 11, 2008; revised February 04, 2009; June 07, 2009; September 06, 2009; and November 02, 2009; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor D. Rubenstein. First published March 25, 2010; current version published August 18, 2010.

S. Sengupta is with Microsoft Research, Redmond, WA 98052 USA (e-mail: sudipta@lcs.mit.edu).

S. Rayanchu and S. Banerjee are with the University of Wisconsin–Madison, Madison, WI 53706 USA (e-mail: shravan@cs.wisc.edu; suman@cs.wisc.edu).

Digital Object Identifier 10.1109/TNET.2010.2042727

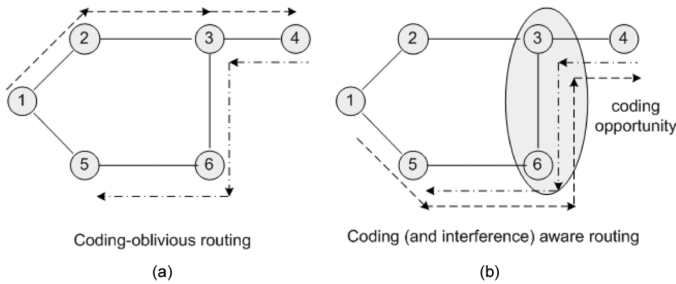


Fig. 2. Coding-oblivious versus coding-aware routing.

(e.g., based on shortest-hop routing, ETX [3], and WCETT [5]), what is the potential benefit of using network coding on end-to-end throughput? Note that this question assumes that the routing decisions are made by an independent routing protocol and our goal is to quantify bounds on throughput gains of network coding under this scenario.

**Coding-Aware (and Interference-Aware) Routing:** Prior work has shown that wireless routing protocols need interference (and link-quality) awareness for improved performance [3], [5], [9]. Introduction of network coding, therefore, raises new questions in the route selection process—should routing decisions now need to be aware of coding opportunities, and if so, what are the approaches to design coding-aware routing in wireless networks?

We illustrate the new opportunity presented by network coding in the context of route selection using a simple example (Fig. 2). The example shows two flows, one from node 1 to node 4 and the other from node 4 to node 5. The link transmission rates are set to 1 unit, and the value of each flow is also set to 1 unit. If we assume a simple scenario where there are no losses on these wireless links, then Fig. 2(a) shows the best paths for the two flows in absence of network coding. These are the shortest and minimum interference paths for the flows, which result in an end-to-end throughput of 0.25. However, if the nodes are allowed to perform network coding, then the throughput of these flows can be improved by choosing paths for the two flows as shown in Fig. 2(b). Note that such a choice increases the path overlap of the two flows to increase coding opportunities. Using the techniques developed in this paper, it can be shown that routing the flows as in Fig. 2(b) results in a throughput of 0.3325, an improvement of 33% compared to the previous case.

Clearly, there is a tradeoff between routing choices that facilitate more coding and routing choices that mitigate interference in the network. We, therefore, present a systematic approach for choosing routes that optimize the tradeoffs between the conflicting effects of increased coding opportunities and increased wireless interference.

**Impact of Multipath Routing:** Many network traffic engineering techniques have shown that multipath routing approaches are known to better utilize the capacity of any network. Therefore, in this paper, we consider coding aware multipath routing. More specifically, we study how the ability to route traffic demands along multiple paths is enhanced by the ability to code packets inside the wireless network. Our formulations can also be used in single-path routing scenarios

where a path needs to be determined in a coding-aware manner or is computed based on a coding-oblivious metric like ETX.

We answer all of these questions in the context of an opportunistic network coding scheme such as COPE [10]. COPE uses the easy-to-implement and relatively inexpensive XOR operation to perform coding. In addition, COPE’s approach to network coding has two other attractive properties.

- 1) *Opportunistic Coding:* Each wireless node uses *only* packets in its local queues for coding (the rules are described in Section II-B). This allows benefits of network coding through local decisions without requiring any form of global coordination between different nodes.
- 2) *Opportunistic Listening:* Exploiting the broadcast nature of the wireless medium, COPE sets each node into a promiscuous mode to snoop on all packets communicated by its neighbors. The snooped packets are used in coding decisions.

We illustrate the advantage of opportunistic listening using the example in Fig. 1(b), where there are four intended packet transfers as follows: from 1 to 3, 3 to 1, 4 to 5, and 5 to 4. Due to range limitations, all transfers need to go via node 2. Let us assume that nodes 1, 3, 4, and 5 transmit their packets in sequence to packet 2. When node 1 transmits its packet to node 2, nodes 4 and 5, in promiscuous mode, snoop on the packet. Similarly, when node 4 transmits its packet to 2, nodes 1 and 3 snoop on this packet. Therefore, at the end of these four packet transmissions, if node 2 were to transmit a single coded packet that XORs all of the four packets, then each node (1, 3, 4, and 5) would be able to correctly decode their intended packets. Thus, the packet transfers are completed by using just five packet transmissions. Note that in absence of coding, eight packet transmissions would have been necessary, while coding without opportunistic listening would have required six.

#### A. Prior Work on Network Coding

The notion of network coding was first proposed by Ahlswede *et al.* [1] for achieving the min-cut capacity in the context of multicast communication. Since then, a large body of work has explored efficient construction of network codes, e.g., [2], [4], [11], [15], and [18]. In [22], the authors provide a linear optimization-based approach for network code construction for multiple unicasts. In the context of wireless networks, Lun *et al.* [16], [17] studied the problem of minimum-cost (energy) multicast involving a single session with a single source node. Ramamoorthy *et al.* [19] derived results for maximum flow achievable in random wireless networks (modeled as geometric random graphs) for a similar single multicast session with a single source.

Li *et al.* [13], [14] show that in some multihop wireless scenarios with multiple unicast sessions, network coding would provide marginal benefits over traditional approaches that do not involve network coding. Ho *et al.* [8] consider network coding across multiple unicasts within the class of network codes *restricted to XOR coding between pairs of flows*. To the best of our knowledge, there is no prior work analyzing the benefits of COPE-type opportunistic network coding for unicast traffic in wireless networks or making routing decisions aware of coding opportunities.

A dynamic routing-scheduling-coding strategy for wireline networks for serving multiple unicast sessions when linear network coding is allowed across sessions is proposed in [26]. Our work differs in that it applies to wireless networks with broadcast capability, models the details of the COPE opportunistic network coding protocol [10], and adds coding-aware routing on top of it.

### B. Unique Contributions of Our Work

Ours is the first work that provides a detailed analytical evaluation of a practical network coding approach, such as COPE, that is applicable to wireless environments with multiple unicast sessions. More specifically, our solutions are applicable to *any* multihop wireless network topology and *any* pattern of concurrent unicast traffic demands. Our results are valid both in presence and absence of opportunistic listening mechanisms. In contrast, the COPE paper [10] constructs and analyzes the best-case bounds for reduction in number of transmissions with opportunistic network coding.

A second important contribution is that this paper introduces the notion of joint coding-aware and interference-aware routing in multihop wireless networks. It illustrates the tradeoffs between needs of increased coding and decreased interference in a systematic manner and identifies efficient routing choices by selecting the appropriate operating point. (Note that COPE does not consider coding-aware routing.)

Finally, this paper illustrates how a coding approach such as COPE can be integrated with a multipath routing solution to further increase end-to-end throughput.

The difficulty of the optimization problem tackled in this paper arises from at least two aspects. First, for a given routing, many combinations of coding opportunities at different nodes are possible, and a subset needs to be selected from the available ones so as to optimize a global objective (e.g., network throughput). Second, when the routing is made aware of the coding opportunities, it has to make choices between routing flows “close to each other” for utilizing coding opportunities and “away from each other” for avoiding interference.

Note that we do not define a full-fledged network coding protocol in this paper, but instead focus on algorithmic analysis (using linear programming-based formulations) that quantifies potential benefits across arbitrary wireless topologies, demands, as well as impact of joint network coding and interference-aware routing techniques. Our framework and its evaluation are fairly general—they properly model arbitrary interference between wireless nodes, availability of different data rates, link loss rates, and other usual practical phenomenon observed in wireless environments. We believe our work provides interesting insights to design protocols that integrate network coding and routing selection techniques.

Independently, and in the same conference as [20], the authors in [23] propose a link cost metric for routing in multihop wireless networks that can utilize network coding opportunities. Their work can be viewed as complementary to ours and can be used for routing connections one at a time (as they arrive over time) (i.e., in an online setting) in a coding-aware manner.

### C. Roadmap

The rest of the paper is structured as follows. In the next section, we introduce notation and formally state the rules for (COPE-type) network coding and our modeling assumptions. In Section III, we consider how to schedule broadcast transmissions for network coding subject to wireless transmit/receive diversity and link interference. In Section IV, we consider maximum throughput coding-aware routing without opportunistic listening and give a linear programming formulation for the multipath routing version of the problem. In Section V, we add opportunistic listening to our optimization framework. In Section VII, using the theoretical formulations developed, we evaluate the benefits of network coding (with and without coding-aware routing) over traditional routing (without coding) on various topologies, including some derived from real mesh network testbed deployments.

## II. NETWORK CODING: NOTATION AND MODELING ASSUMPTIONS

### A. Notation

The wireless network topology, given by the nodes and the links corresponding to pairs of nodes within direct communication range, is modeled as a graph  $G = (N, E)$  with node set  $N$  and (directed) edge set  $E$ . Each node in the network can be a source or destination of traffic. The sets of incoming and outgoing edges at node  $i$  are denoted by  $E^-(i)$  and  $E^+(i)$ , respectively. We let  $e = (i, j)$  represent a directed link in the network from node  $i$  to node  $j$ . The transmitting node for link  $e$  will be denoted by  $t(e)$  and its receiving node by  $r(e)$ . We will denote the reverse of link  $e = (i, j)$  by  $\bar{e} = (j, i)$ . The rate of transmission on link  $e$  will be denoted by  $R_e$  and its delivery probability by  $p_e$ . Thus, the *effective rate of transmission* on link  $e$  is  $u_e = p_e R_e$ .

Let  $D$  be the set of demands. A demand  $k \in D$  has source node  $s(k)$ , destination node  $d(k)$ , and traffic value  $t(k)$ . For a given routing/coding scheme, the *throughput* is defined as the maximum multiplier  $\lambda$  such that all demands with their traffic values multiplied by  $\lambda$  can be feasibly routed by the network. In this paper, we will be concerned with maximizing the throughput  $\lambda$  for coding-aware network routing.

For a path  $P$  and links  $e, e_1, e_2$ , we will use  $e \in P$  to denote that link  $e$  is on path  $P$  and  $e_1 e_2 \in P$  to denote that path  $P$  contains link  $e_1$  followed by link  $e_2$  in *consecutive order* (this assumes  $r(e_1) = t(e_2)$ ). For a path  $P$  and node  $i$ , we will use  $i \in P$  to denote that node  $i$  is on path  $P$ .

### B. Coding Rules and Modeling Assumptions

Consider  $k$  packets  $p_1, p_2, \dots, p_k$  at a node that have distinct next-hop nodes  $n_1, n_2, \dots, n_k$ , respectively. Suppose these are coded together to form the coded packet  $p = p_1 \oplus p_2 \oplus \dots \oplus p_k$  that is broadcast to all the above next-hop nodes. This is a valid network coding if the next-hop node  $n_i$  for each packet  $p_i$  already has all other packets  $p_j$  for  $j \neq i$  (so that it can decode  $p_i$ ). This can happen if:

- 1) node  $n_i$  is the previous-hop node of packet  $p_j$ ; or
- 2) node  $n_i$  overheard packet  $p_j$  from the transmission of its previous-hop node (opportunistic listening).

In the specification of coding opportunities, node  $n_i$  can have packet  $p_j$  also if: 1') node  $n_i$  was one of the nodes traversed by packet  $p_j$  before its previous-hop node; or 2') node  $n_i$  overheard packet  $p_j$  from the transmission of a node traversed by packet  $p_j$  before its previous-hop node. For either of conditions 1') or 2') to hold, node  $n_i$  will need to buffer packet  $p_j$  for longer periods of time in the hope that additional coding opportunities involving packet  $p_j$  will arise at a node farther downstream on its path (and adjacent to  $n_i$ ). (It is also unlikely that packet  $p_j$  will take a circuitous route and pass through a neighbor node of  $n_i$  at least two-hops downstream when it has already passed through either  $n_i$  or one of its neighbors.) In contrast, conditions 1) and 2) do not require node  $n_j$  to buffer packet  $p_j$  beyond the transmission at the next-hop node after the packet passed through or was overheard.

In this paper, we do not model coding opportunities arising from conditions 1') and 2') above. Our assumption is consistent with the requirement to have minimal additional packet buffering requirements at each wireless node for network coding.

### III. SCHEDULING BROADCAST TRANSMISSIONS

Network coding exploits the broadcast nature of the wireless medium, hence a proper model that handles scheduling of broadcast transmissions is essential. Accordingly, prior to developing our analytical framework for network coding, we discuss how to handle broadcast transmissions in this section. We use the protocol model of interference introduced by Gupta and Kumar [7]. Prior work by Jain *et al.* [9] developed an optimization framework for scheduling *unicast* transmissions. We generalize their approach to handle broadcast transmissions. To the best of our knowledge, this generalization has not been proposed earlier in the literature.

#### A. Prior Work on Scheduling Unicast Transmissions

Let  $d_{ij}$  denote the distance between nodes  $i$  and  $j$ . Let the radio at node  $i$  have a communication range of  $\ell_i$  and potentially larger interference range  $\ell'_i$ . Under the *protocol model of interference* [7], if there is a single wireless channel, a transmission from node  $i$  to node  $j$  is successful if: 1)  $d_{ij} \leq \ell_i$  (receiver is within communication range of sender); and 2) any node  $k$ , such that  $d_{kj} \leq \ell'_k$ , is not transmitting (receiver is free of interference from any other possible sender).

The authors in [9] construct a conflict graph whose nodes correspond to links in the topology graph. Two nodes are connected by an (undirected) edge in the conflict graph if the corresponding links cannot be scheduled simultaneously. Scheduling link transmissions for transmit/receive diversity and link interference are then modeled using constraints corresponding to *either* cliques or independent sets in the conflict graph. It is shown that the clique constraints provide an upper bound on the throughput (and this upper bound is not always tight), while the independent constraints provide an achievable lower bound. Note that this does not completely model the IEEE 802.11 MAC [6], primarily because it does not take contention time for acquiring a channel into account.

#### B. Broadcast Transmission Rates

We are given the rate of transmission  $u_e$  and associated delivery probability  $p_e$  for each wireless link  $e$ . The delivery probability of a wireless link generally decreases with increase in transmission rate. For a given rate of transmission, the delivery probability is most accurately determined only by experiment, as it depends on many factors, including environment, transmission power, distance between transmitter and receiver, channel fading, and background noise. We discuss how we use given information on a single  $(u_e, p_e)$  per link  $e$  to obtain broadcast transmission rates for our model.

Let  $B$  be a subset of outgoing links at some node. If  $B = \{e\}$  consists of a single link, then its effective rate of transmission is simply  $u(B) = u_e$ . Now, consider the case when  $B$  consists of multiple links. Because of the given limited information per link (as explained above), we will assume that the transmission rate for broadcast on  $B$  is the minimum rate of its component links, which equals  $R(B) = \min_{e \in B} R_e$  (this allows us to make a conservative estimation of the effective rates of transmission for broadcasts). To compute the effective rate of transmission, we need the delivery probability of this transmission. Since  $R(B) \leq R_e$  for any  $e \in B$ , the probability that a single broadcast reaches  $r(e)$  is *at least*  $p_e$  since delivery probabilities can only increase with decrease in transmission rates. Assuming that losses on individual links are independent, the delivery probability of the broadcast  $B$  is at least  $\prod_{e \in B} p_e$ . We will use this conservative estimate as the delivery probability for broadcast on  $B$ . Thus, the effective rate of transmission for broadcast on  $B$  is given by

$$u(B) = \left( \prod_{e \in B} p_e \right) \min_{e \in B} R_e.$$

*Note that the optimization techniques developed in this paper are independent of the manner in which effective broadcast transmission rates are obtained.*

#### C. Broadcast Conflict Graph

A broadcast transmission at node  $i$  on a subset  $B$  of its outgoing links will be represented as  $(i, B)$  and the associated broadcast traffic as  $q_i^B$ . Note that this includes unicast as a special case when the set  $B$  consists of just one link. We define the *broadcast conflict graph*  $F$  as a natural extension of the conflict graph for unicast transmissions. Each node in this graph represents a broadcast transmission  $(i, B)$ . Let  $r(B)$  denote the set of receiver nodes for the links in broadcast set  $B$ . Two broadcasts  $(i_1, B_1)$  and  $(i_2, B_2)$  interfere and hence have an edge between them in the broadcast conflict graph if either:

- some node  $j \in r(B_1)$  is within interference range of node  $i_2$ ; or
- some node  $j \in r(B_2)$  is within interference range of node  $i_1$ .

It can be verified that the above conditions include special cases like  $i_1 = i_2$  (broadcasts out of same node) or  $r(B_1) \cap r(B_2) \neq \Phi$  (broadcasts have a common receiver). We note that a similar model capturing the conflict between a set of broadcast transmissions has been presented subsequently in [21] and [24] after our work [20] was published. In these later papers, the notion of

concurrent transmitter sets, denoting the set of broadcast transmissions that are simultaneously active, is used to model the conflicts.

The idea of a conflict graph has been used to model constraints in network code construction [27], [28]. In this paper, we use a broadcast conflict graph to schedule wireless broadcast transmissions, not for network code construction itself.

#### D. Clique Constraints for Broadcast Transmission Scheduling

With this generalization of the conflict graph, we can use constraints corresponding to cliques in the broadcast conflict graph. Consider a *clique* in the broadcast conflict graph. Let  $C$  be the set of broadcast nodes  $(i, B)$  that correspond to nodes of this clique. The fraction of time that broadcast  $(i, B)$  is active is  $y_i^B/u(B)$ . Since the broadcasts in  $C$  mutually conflict with each other, at most one of them can be active at any given time. This can be modeled by the constraint

$$\sum_{(i,B) \in C} \frac{y_i^B}{u(B)} \leq 1 \quad \forall \text{ cliques } C \text{ in } F. \quad (1)$$

Note that it is sufficient to add just constraints corresponding to *maximal cliques* since the constraints corresponding to cliques contained inside maximal cliques are redundant. The clique constraints provide a necessary condition for the rates  $y_i^B$  to be achievable. However, they may not be always be sufficient.

#### E. Independent Set Constraints for Broadcast Transmission Scheduling

Let  $I_1, I_2, \dots, I_q$  denote all the *maximal* independent sets in the broadcast conflict graph (let each set consist of corresponding broadcasts of the form  $(i, B)$ ). Let independent set  $I_j$  be active for  $b_j$  fraction of the time. Any set of active broadcast transmissions are contained in some independent set. Then, we should have

$$\sum_{j=1}^q b_j \leq 1. \quad (2)$$

The fraction of time that an individual broadcast transmission is active is at most the sum of the fraction of time that each independent set it belongs to is active. This can be written as

$$\frac{y_i^B}{u(B)} \leq \sum_{j: I_j \ni (i,B)} b_j \quad \forall \text{ broadcasts } (i, B). \quad (3)$$

When all maximal independent sets are considered, these constraints provide a necessary and sufficient condition for the rates  $y_i^B$  to be achievable. In practice, to keep computation times reasonable, only a subset of maximal independent sets can be considered, in which case the rates that are feasible form a subset of all achievable rates (and hence give a lower bound on the maximum throughput).

We will use the clique constraints in our linear programming formulations for routing with network coding in Sections IV and V. One may choose to use the independent set constraints instead. We do talk about convergence of the computed upper and lower bounds using each set of constraints in the evaluation section.

## IV. CODING-AWARE NETWORK ROUTING WITHOUT OPPORTUNISTIC LISTENING

In the absence of opportunistic listening, a coding opportunity at a node involves XOR-ing exactly two packets. These packets enter and leave the node using the same links but in opposite directions. We summarize this in the following lemma.

*Lemma 1:* Under the assumptions for network coding outlined in Section II-B and in the absence of opportunistic listening, a coding opportunity at a node involves XOR-ing exactly two packets (and not more).

*Proof:* We prove by contradiction. Consider three (or more) packets  $p_1, p_2, p_3$  at a node with distinct next-hop nodes  $n_1, n_2, n_3$ , respectively. For these three packets to be coded together, each next-hop node must be able to decode its respective packet. For node  $n_1$  to be able to decode packet  $p_1$ , it must already have packet  $p_3$ . Because there is no opportunistic listening, the previous-hop node of packet  $p_3$  must be  $n_1$ . Similarly, for node  $n_2$  to be able to decode packet  $p_2$ , it must already have packet  $p_3$ . Hence, the previous-hop node of packet  $p_3$  must be  $n_2$ . This is a contradiction since nodes  $n_1$  and  $n_2$  are distinct. ■

We now formulate the problem of maximum throughput routing with network coding without opportunistic listening. We consider multipath routing. (By using integer variables, the formulation can be modified to handle single-path routing.) Let  $\mathcal{P}^k$  denote the set of available paths from source  $s(k)$  to destination  $d(k)$  for routing demand  $k$ . For example, we could choose the  $K$ -shortest distance paths from  $s(k)$  to  $d(k)$  as the set  $\mathcal{P}^k$ . Let routing variable  $f^k(P)$  denote the amount of traffic on path  $P$  for routing demand  $k$ . Let variable  $y_i^B$  denote the traffic that is broadcast at node  $i$  on link set  $B \subseteq E^+(i)$ . Because there is no opportunistic listening, we have  $|B| \leq 2$ . If  $|B| = 1$ , the corresponding transmission is a unicast (of a native packet) on the single link in  $B$ . Let  $\lambda$  denote the throughput for routing all demands in  $D$ . Then, the problem of routing under network coding without opportunistic listening so as to maximize throughput can be expressed as the following linear program (LP):

$$\begin{aligned} & \text{maximize } \lambda \\ & \text{subject to} \end{aligned} \quad (4)$$

$$\sum_{P \in \mathcal{P}^k} f^k(P) = t(k)\lambda \quad \forall k \in D \quad (5)$$

$$\begin{aligned} y_i^{\{e_1, e_2\}} & \leq \sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni \bar{e}_1 \bar{e}_2} f^k(P) \\ & \quad \forall e_1, e_2 \in E^+(i), i \in N \end{aligned} \quad (6)$$

$$\begin{aligned} y_i^{\{e_1, e_2\}} & \leq \sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni \bar{e}_2 \bar{e}_1} f^k(P) \\ & \quad \forall e_1, e_2 \in E^+(i), i \in N \end{aligned} \quad (7)$$

$$y_i^{\{e\}} = \sum_{k \in D, s(k)=i} \sum_{P \in \mathcal{P}^k, P \ni e} f^k(P)$$

$$\begin{aligned} & + \sum_{e_1 \in E^-(i)} \left[ \sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni e_1 e} f^k(P) \right. \\ & \quad \left. - y_i^{\{e, \bar{e}_1\}} \right] \\ & \quad \forall e \in E^+(i), i \in N \end{aligned} \quad (8)$$

$$\sum_{(i,B) \in C, |B| \leq 2} \frac{y_i^B}{u(B)} \leq 1 \quad \forall \text{ cliques } C \text{ in } F. \quad (9)$$

Constraints (4) state that the total traffic routed on the available paths for a demand must equal the demand value multiplied by its throughput. Constraints (5) and (6) determine the maximum amount of coded traffic  $y_i^{\{e_1, e_2\}}$  that can be broadcast on outgoing links  $\{e_1, e_2\}$  at node  $i$ . The total traffic traversing node  $i$  along link sequence  $\bar{e}_1 e_2$  is  $\sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni \bar{e}_1 e_2} f^k(P)$  and along link sequence  $\bar{e}_2 e_1$  is  $\sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni \bar{e}_2 e_1} f^k(P)$ . Thus,  $y_i^{\{e_1, e_2\}}$  is at most each of these amounts.

Constraint (7) gives the total amount of traffic  $y_i^{\{e\}}$  that is unicast on outgoing link  $e$  at node  $i$ . This traffic can be divided into two parts. The first part is the traffic that originates at node  $i$  and is sent on link  $e$  and equals  $\sum_{k \in D, s(k)=i} \sum_{P \in \mathcal{P}^k, P \ni e} f^k(P)$ . The second part is the amount of transit traffic at node  $i$  with next-hop  $e$  that could not be coded with other flows and equals  $\sum_{e_1 \in E^-(i)} [\sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni e_1 e} f^k(P) - y_i^{\{e, \bar{e}_1\}}]$ .

Finally, constraints (8) are the broadcast transmission scheduling constraints corresponding to cliques in the broadcast conflict graph, as discussed in Section III. The cliques are restricted to include broadcast sets of size at most 2 since there is no opportunistic listening.

Note that this LP can also be used to compute the throughput when ETX metric routes are used. For this purpose, the set  $\mathcal{P}^k$  of available paths for demand  $k$  consists of the (singleton) shortest-cost ETX path from node  $s(k)$  to node  $d(k)$ .

The path-indexed routing variables  $f_k(P)$  can be reduced to polynomial size by converting to dual-link-indexed variables  $f_k(e_1, e_2)$ , where  $(e_1, e_2)$  denotes the incoming–outgoing link pair at each node. This corresponds to routing on a graph with a (straightforward) node-splitting transformation, which we will describe. For each other node  $i \in N$  with (incoming/outgoing) degree  $d_i$  in the original graph, create nodes  $a_1^i, a_2^i, \dots, a_{d_i}^i$  and  $b_1^i, b_2^i, \dots, b_{d_i}^i$  in the new graph. Assume that the incoming and outgoing edge pairs at each node in the original graph are ordered in some sequence. For each combination  $k \neq \ell$  of edge pair numbers at node  $i$ , the  $k$ th incoming edge—say  $e_1$ —at node  $i$  in the original graph enters node  $a_k^i$  in the new graph. Likewise, the  $\ell$ th outgoing edge—say  $e_2$ —at node  $i$  in the original graph exits at node  $b_\ell^i$  in the new graph. Also, there is a directed edge from node  $a_k^i$  to node  $b_\ell^i$  corresponding to the dual-link  $(e_1, e_2)$ . The amount of flow on this (dual-link) edge in the transformed graph for demand  $k$  now corresponds to the variable  $f_k(e_1, e_2)$ .

## V. CODING-AWARE NETWORK ROUTING WITH OPPORTUNISTIC LISTENING

In this section, we develop a model that allows coding of different packets at a node to take advantage of opportunistic listening by its neighbors of the transmission of packets by the previous-hop node of each packet. We represent a coding opportunity by what we call a *coding structure*. We give a linear programming formulation for maximum throughput coding-aware routing with opportunistic listening.

### A. Modeling Network Coding With Opportunistic Listening

The usefulness of opportunistic listening depends on whether the listening involved transmission of a coded or native packet.

If a packet was transmitted as a coded packet (i.e., XOR-ed with other packets), then a listening node (that is not its next-hop) will not be able to decode it if it does not have all the other packets. This is likely to be the case more often than not since coding guarantees that only the next-hop node of a packet (and not opportunistic listeners) will be able to decode it. Hence, we will model *useful* listening opportunities that involve transmission of a native packet only. Thus, the coding opportunity for a packet at a node is determined by two factors, namely:

- the combination of its incoming and outgoing links at that node; and
- whether the packet was received at that node as a coded or native packet.

Let us fix a node  $i \in N$  of (in/out) degree  $d_i$ . The number of possible combinations of incoming/outgoing links at this node is  $d_i(d_i - 1)$ . A coding opportunity at node  $i$  is completely specified by a structure  $S$  that consists of elements of the form  $s = (e_1 e_2, v)$ , where  $e_1$  is the incoming link of the packet,  $e_2$  is the outgoing link of the packet, and  $v = c, n$  depending on whether the packet was received as coded ( $c$ ) or native ( $n$ ). We will call such structures *coding structures*. We will denote the  $v$ -component of  $s$  by  $v(s)$ . The previous-hop node of  $s$  is  $t(e_1)$ , and the next-hop node of  $s$  is  $r(e_2)$ .

A coding structure represents a coding opportunity under the following conditions.

- 1) The next-hop node of each  $s \in S$  must be distinct since two packets going to the same next-hop node cannot be coded. Thus, a coding structure  $S$  at node  $i$  can contain at most  $d_i$  elements  $f$ .
- 2) The next-hop node for each  $s \in S$  must be able to decode its packet. Thus, for any given  $s \in S$ , the next-hop node of  $s$  should already have the packet associated with all other  $s' \in S$ ,  $s' \neq s$ . For the latter to hold, we must have for each other  $s' \in S$ , the next-hop node of  $s$  must be either: a) the previous-hop node of  $s'$ ; or b) be a neighbor of previous-hop node of  $s'$  to have overheard the transmission, provided  $v(s') = n$ , i.e., the associated packet was transmitted as native. Note that in case a), the value of  $v(s')$  does not matter.

The broadcast set  $b(S)$  for a coding structure is defined as the set of next-hop links of each  $s \in S$ . This is simply the set of links over which the XOR-ed packets are broadcast for the coding opportunity represented by the structure  $S$ . We will say that  $s$  (or the associated packet) participates in coding structure  $S$  as *native-received* if  $v(s) = n$  and as *coded-received* if  $v(s) = c$ .

We now provide a few simple examples to explain how we capture useful listening opportunities with coding structures. Consider network (a) in Fig. 3 where packets  $p_1, p_2$  have to be sent along the paths  $1 \rightarrow 2 \rightarrow 3$  and  $3 \rightarrow 2 \rightarrow 1$ , respectively. A coding opportunity arises at node 2 as it can forward both the packets in a single transmission  $p = p_1 \oplus p_2$ . Since node 1 is the previous hop of  $p_1$ , it can correctly decode the packet  $p_2$  using  $p \oplus p_1$ , i.e., the ability to decode the packet does not depend on whether the  $p_1$  was native-received or coded-received. Similarly, transmission of packet  $p_2$  from node 3 can be as a native or coded packet. Therefore, the set of coding structures for node 2 is  $S = \{(e_1 e_2, n), (\bar{e}_2 \bar{e}_1, n)\}, \{(e_1 e_2, c), (\bar{e}_2 \bar{e}_1, n)\}, \{(e_1 e_2, n), (\bar{e}_2 \bar{e}_1, c)\}, \{(e_1 e_2, c), (\bar{e}_2 \bar{e}_1, c)\}$ .

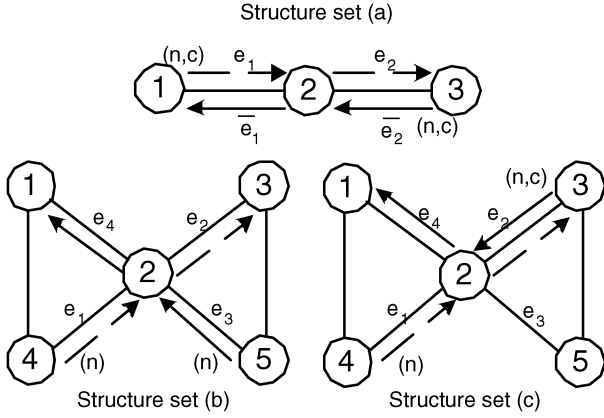


Fig. 3. Network coding structures.

Now, consider the network (b) where packets  $p_1, p_2$  have to be sent along the paths  $4 \rightarrow 2 \rightarrow 3$  and  $5 \rightarrow 2 \rightarrow 1$ , respectively. Assume that node 4 sends the packet  $p_1$  as a native packet. In such a case, node 1 would also be able to receive it. Similarly, node 3 will also receive the packet  $p_2$  if node 5 transmits it as a native packet. Node 2 can then transmit the coded packet  $p = p_1 \oplus p_2$ , and nodes 1 and 3 will be able to correctly decode  $p_2$  and  $p_1$ , respectively. However, such a coding opportunity cannot arise if either of the transmissions  $4 \rightarrow 2$  or  $5 \rightarrow 2$  were coded-received. Hence, only  $S = \{(e_1e_2, n), (e_3e_4, n)\}$  would represent a coding opportunity for node 2. It is easy to see that for network (c), the structures  $\{(e_1e_2, n), (e_3e_4, n)\}$  and  $\{(e_1e_2, n), (e_3e_4, c)\}$  would represent coding opportunities.

We use a straightforward method to generate all the valid coding structures. Let  $\Gamma_i$  denote the set of all valid coding structures at node  $i$ . The number of different elements  $s = (e_1e_2, v)$  at node  $i$  is  $2d_i(d_i - 1)$ . Since nodes in a wireless mesh network topology have small degree, this number is small. Hence, it is relatively fast to generate all such combinations  $S$  of elements  $s$  obeying the coding opportunity condition 1) and checking whether condition 2) holds.

For structures containing three or more elements, we can further reduce the number of candidates by using the following lemma.

**Lemma 2:** For a coding structure  $S$  of size at least 3 ( $|S| \geq 3$ ) that represents a valid coding opportunity, all elements  $s \in S$  must have  $v(s) = n$ . That is, the packets involved in this coding opportunity must have been received as native packets at the respective node.

It is important to note that the size and running time of the LP formulation for coding-aware network routing with opportunistic listening in Section V is independent of the method used to generate the structures (but depends on the number of structures in  $\Gamma_i$ ).

We explain why we need not consider additional transmission scheduling constraints for opportunistic listening. Consider a node  $n_1$  that listens on the transmission of a (native) packet from a node  $n_2$  to its next-hop node  $n_3$ . We argue that node  $n_1$  cannot be transmitting or receiving from another node during this time in any valid schedule. Since  $n_1$  is a neighbor of  $n_2$  and hence within interference range of  $n_2$ , it cannot be a receiver for some other transmission during this time. Also, for the listening to be

useful in a coding decision at  $n_3$ ,  $n_1$  must be a neighbor of  $n_3$ . Thus,  $n_1$  cannot also be transmitting during this time, otherwise it will interfere at  $n_3$  with the transmission from  $n_2$ .

## B. Maximizing Throughput

$$\begin{aligned}
& \text{maximize } \lambda \\
& \text{subject to} \\
& \sum_{P \in \mathcal{P}^k} f^k(P) = t(k)\lambda \quad \forall k \in D \quad (9) \\
& \sum_{S \ni (e_1e_2, n)} x_i(S) \leq \sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni e_1e_2} z_{t(e_1)}^k(P) \\
& \quad \forall e_1 \in E^-(i), e_2 \in E^+(i), \\
& \quad i \in N \quad (10) \\
& \sum_{S \ni (e_1e_2, c)} x_i(S) \leq \sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni e_1e_2} [f^k(P) \\
& \quad - z_{t(e_1)}^k(P)] \\
& \quad \forall e_1 \in E^-(i), e_2 \in E^+(i), \\
& \quad i \in N \quad (11) \\
& \sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni e_1e_2} f^k(P) = \sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni e_1e_2} z_i^k(P) \\
& \quad + \sum_{S \ni (e_1e_2, n)} x_i(S) \\
& \quad + \sum_{S \ni (e_1e_2, c)} x_i(S) \\
& \quad \forall e_1 \in E^-(i), e_2 \in E^+(i), \\
& \quad i \in N \quad (12) \\
& z_{s(k)}^k(P) = f^k(P) \quad \forall P \in \mathcal{P}^k, k \in D \quad (13) \\
& z_i^k(P) \leq f^k(P) \quad \forall i \in P - \{s(k), d(k)\}, \\
& \quad P \in \mathcal{P}^k, k \in D \quad (14) \\
& y_i^{\{e\}} = \sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni e} z_i^k(P) \\
& \quad \forall e \in E^+(i), i \in N \quad (15) \\
& y_i^B = \sum_{S \in \Gamma_i, b(S)=B} x_i(S) \\
& \quad \forall i \in N, |B| \geq 2 \quad (16) \\
& \sum_{(i,B) \in C} \frac{y_i^B}{u(B)} \leq 1 \quad \forall \text{ cliques } C \text{ in } F. \quad (17)
\end{aligned}$$

In addition to the variables for the first linear programming formulation, we will use two new sets of variables. Let  $x_i(S)$  denote the traffic associated with coding structure  $S$  at node  $i$ —this is the traffic amount associated with each  $e_1e_2$  link-pair participating in the structure. Also, let  $z_i^k(P)$  be the portion of the traffic on path  $P$  for demand  $k$  that is transmitted as native from node  $i$  (the variables  $z_i^k(P)$  are defined for all  $i \in P - \{d(k)\}$ ). Also, the broadcast set  $B$  in variables  $y_i^B$  are no longer restricted to be of size at most 2, but can include all the outgoing links at node  $i$ .

Constraint (10) states that for each combination of incoming link  $e_1$  and outgoing link  $e_2$  at node  $i$ , the portion of transit

traffic that participates in coding as native-received flows is at most the amount that was received as native from  $t(e_1)$ . Similarly, constraint (11) states that the portion of transit traffic that participates in coding as coded-received flows is at most the amount that was received as coded.

Constraint (12) consists of balance constraints for the total transit traffic entering through link  $e_1$  and exiting through link  $e_2$  at node  $i$ , which appears on the left-hand side (LHS). The first portion on the right-hand side (RHS), namely,  $\sum_{k \in D} \sum_{P \in \mathcal{P}^k, P \ni e_1, e_2} z_i^k(P)$ , is the amount of transit traffic that goes out as native (i.e., does not participate in any coding). The second portion,  $\sum_{S \ni (e_1, e_2, n)} x_i(S)$ , is the amount of transit traffic that participates in coding as native-received flows. The third portion,  $\sum_{S \ni (e_1, e_2, c)} x_i(S)$ , is the amount of transit traffic that participates in coding as coded-received flows.

Constraints (13) and (14) are the boundary conditions for the  $z_i^k(P)$  variables. Constraint (13) states that the source node of every path transmits the entire traffic on that path as native since no coding opportunities are available for originating traffic at the source node. Constraint (14) states that, for a given path, the amount of traffic transmitted as native at each transit node is at most the total traffic on that path.

Constraint (15) expresses the unicast traffic variable  $y_i^{\{e\}}$  as the total amount of traffic that is transmitted as native on link  $e$  at node  $i$ . Constraint (16) expresses the broadcast traffic variable  $y_i^B$  as the total amount of traffic that is transmitted as coded on link set  $B$  at node  $i$ ; this corresponds to the sum of traffic over all coding structures  $S \in \Gamma_i$  with  $b(S) = B$ .

Finally, as before, constraint (17) consists of the broadcast transmission scheduling constraints corresponding to cliques in the broadcast conflict graph. In this case, the broadcast sets of the cliques are not restricted to be of size 2 since opportunistic listening is allowed.

In a manner analogous to Section IV, this LP can also be used to compute the throughput when ETX metric routes are used. For this purpose, the set  $\mathcal{P}^k$  of available paths for demand  $k$  consists of the (singleton) shortest-cost ETX path from node  $s(k)$  to node  $d(k)$ .

## VI. SOME EXTENSIONS TO OUR PROBLEM FORMULATIONS

In this section, we consider two extensions to our earlier formulations. The first concerns minimizing total transmission power for network coding in wireless sensor networks where the nodes are limited by battery power. The second extension concerns restricting the routing for each demand to be along a single path (instead of multipath).

**Minimizing Total Transmission Bandwidth in Wireless Sensor Networks:** Conserving battery power is an important optimization goal in sensor networks. In this case, minimizing total transmission bandwidth may be a more important metric than maximizing throughput. To handle this in our earlier linear programming formulations, we set the throughput  $\lambda = 1$  and change the objective function to

$$\text{minimize } \sum_{(i,B)} y_i^B.$$

**Single-Path Routing:** The linear programming formulations in Sections IV and V can be modified to route demands along single paths (instead of splitting traffic across multiple paths) as follows.

- The routing variables  $f^k(P)$  now become 0/1 variables. The value of  $f^k(P)$  is 1 if the demand  $k$  is routed along path  $P$ , and is 0 otherwise.
- The first constraint in each LP is replaced by

$$\sum_{P \in \mathcal{P}^k} f^k(P) = 1 \quad \forall \quad k \in D$$

which states that exactly one path is chosen for routing each demand.

- In all other constraints where  $f^k(P)$  appears, it is replaced by  $t(k)f^k(P)$ , which is now the traffic on path  $P$  if it is (the only path) used for demand  $k$ .
- Since scaling the demands by  $\lambda$  increases each  $y_i^B$  by the same factor, the clique constraints for scheduling transmissions now become  $\sum_{(i,B) \in C} y_i^B / u(B) \leq 1/\lambda$ . Using the variable transformation  $\bar{\lambda} = 1/\lambda$ , we write these constraints as

$$\sum_{(i,B) \in C} \frac{y_i^B}{u(B)} \leq \bar{\lambda} \quad \forall \quad \text{cliques } C \text{ in } F.$$

- Maximizing  $\lambda$  is now equivalent to minimizing  $\bar{\lambda}$ . The objective function becomes

$$\text{minimize } \bar{\lambda}.$$

With the introduction of integer 0/1 variables, the formulations now become mixed integer linear programming (MILP) problems. Like linear programs with continuous variables, MILPs can also be solved using the CPLEX [25] package.

## VII. EVALUATION

We evaluate the performance of the following schemes on a variety of network topologies: 1) shortest-path routing (SPATH); 2) interference-aware multipath routing (MPATH) (computed using a variation of the LP formulation in [9]); 3) shortest-path routing with network coding (SPATH-CODE), derived by restricting  $\mathcal{P}^k$  to (singleton) shortest-cost ETX path; 4) **network coding-aware multipath routing** (CA-MPATH-CODE, shortened to LP-CODE in the plots); and 5) **network coding-aware single-path routing** (CA-SPATH-CODE).<sup>2</sup> We evaluate SPATH-CODE and CA-MPATH-CODE both with and without opportunistic listening. The main goals of our simulation are to: 1) characterize the maximum throughput of a given network and traffic patterns with network coding; 2) quantify the benefits of coding aware routing. For the evaluations using multipath routing, the set  $\mathcal{P}_k$  of available paths was chosen to consist of the  $K$ -shortest distance paths for  $K = 40$ . (In all cases, the actual number of paths used by the optimal solution

<sup>2</sup>Due to space constraints, we present results for CA-SPATH-CODE for only one illustrative topology.



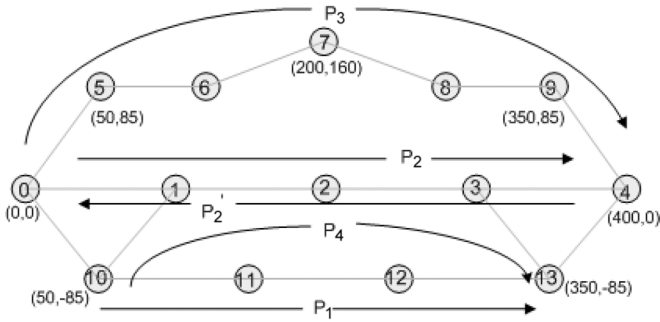


Fig. 4. Example topology.

was much smaller.) We solve all linear programs using CPLEX [25].

#### A. Results on Illustrative Topology

We first start with a simple 14-node topology shown in Fig. 4. We have deliberately chosen this topology to illustrate two key properties of our LP formulations. For a given network and traffic demands, the LP computes the routes to obtain maximum throughput by: 1) maximizing the coding opportunities; and 2) minimizing the interference among the flows.

We assume that the communication range of each node is 100 units and the interference range is 200. For simplicity, assume that all the wireless links have the same transmission rates (1 unit). We assume the protocol interference model (Section III). We allow opportunistic listening. We compare the maximum throughput obtained using CA-SPATH-CODE, SPATH-CODE, and SPATH for the following scenarios.

- **Scenario A—Minimizing interference:** Consider two concurrent demands  $d_1 : 0 \rightarrow 4$  and  $d_2 : 10 \rightarrow 13$ , each of 1 unit, which have to be routed. Now, SPATH and SPATH-CODE will route the demands  $d_1$  and  $d_2$  along the shortest paths  $P_1$  and  $P_2$ , respectively. It can be shown that only one of these seven links can be active at a time, and hence the optimal throughput with SPATH would be  $1/7$ . Since there are no coding opportunities along these paths, the optimal throughput with SPATH-CODE would also be  $1/7$ . However, the routes computed by CA-SPATH-CODE result in an optimal throughput of  $2/9$ , i.e., an improvement of 55% over SPATH-CODE. This is because, in the absence of coding opportunities, CA-SPATH-CODE routes the demands along the paths that result in the least interference. In this case, CA-SPATH-CODE routes  $d_2$  on a longer path  $P_3$ , which yields a better throughput because of the reduced interference with  $d_1$ .
- **Scenario B—Maximizing coding opportunities:** We now reverse the first demand ( $d_1 : 4 \rightarrow 0$ ) and compute the optimal throughput for each scheme. In this case, SPATH and SPATH-CODE route the demand along the paths  $P_1$  and  $P_2'$  (Fig. 4). For these routes, an optimal schedule results in a throughput of  $1/5$  (since links  $\{(10, 11), (4, 3)\}, \{(1, 0), (12, 13)\}$  can be scheduled in single time slot). Routing the demand  $d_1$  along the  $P_3'$  (reverse of  $P_3$ ) does not reduce the interference in this case and results in a throughput of  $1/5$ . However, CA-SPATH-CODE takes advantage of opportunistic listening and routes the demand  $d_1$  along path  $P_4$  and  $d_2$

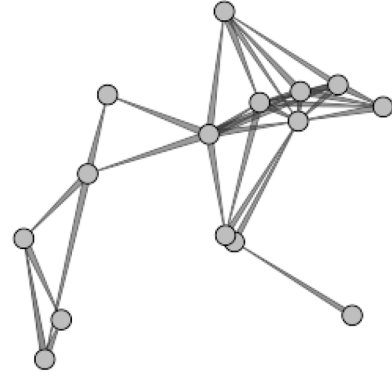


Fig. 5. 15-node random topology.

along  $P_2'$ , which results in coding opportunities at nodes 1, 2, and 3. This reduces the number of transmissions from seven to four and hence results in an optimal throughput of  $1/4$  (an improvement of 25% over SPATH-CODE). It is important to note that a longer path  $P_4$  results in a better throughput because of increased coding opportunities.

#### B. Evaluation on Wireless Network Topologies

We evaluated the performance of our LP formulations on various topologies. Here, we present the results for a random graph topology and two other topologies derived from an operational community wireless network and an in-building mesh network.

**Optimality:** In general, the upper and the lower bounds computed by the clique constraints and the independent set constraints (Section III) may not converge. However, for the topologies under consideration, the upper and lower bounds for each of the schemes were found to converge, hence assuring us of the optimality of the solution.

**Random Graph Topology:** We present the results for a 15-node wireless network shown in Fig. 5. The positions of the nodes were chosen randomly in a square of side 400 units. The communication range of each node is set to 100 units, and interference range is set to 200. The capacity of the communication links is chosen according to Shannon's formula (parameters in the formula are appropriately fixed). The average degree of each node was 4.4. The source and the destination for each flow are chosen at random. We vary the number of demands from 10 to 600.

**Evaluation #1—Interplay of interference effects and coding opportunities:** We study the variation of throughput gain as a function of number of demands for SPATH-CODE, MPATH, and CA-MPATH-CODE. We allow opportunistic listening. Fig. 8 shows the throughput normalized w.r.t. shortest-path routing for increasing number of flows. The throughput gain for MPATH increases with the increase in number of flows. However, as the network load increases, it becomes difficult to find paths that do not interfere with other flows, and thus the gain decreases. For SPATH-CODE, the routes are fixed, hence interference cannot be avoided, and the gains are purely due to coding opportunities. Coding opportunities increase with the number of flows, resulting in increasing throughput gain. We observe that CA-MPATH-CODE outperforms both the schemes with a maximum gain of 65% over shortest path and of about

40% over SPATH-CODE. The higher slope of the curve is due to the combined gains of reduced interference and increased coding opportunities. It is important to note that increase in coding opportunities is due to coding-aware routing as well as opportunistic listening.

**Evaluation #2—Coding-aware routing versus opportunistic listening:** A natural question is to ask how much is the performance improvement due to coding aware routing (without opportunistic listening) alone and how much is due to opportunistic listening. Motivated by this question, we study the variation of throughput gain as a function of increasing number of demands for SPATH-CODE and CA-MPATH-CODE, both with and without opportunistic listening. The results are shown in Fig. 9. Allowing opportunistic listening for SPATH-CODE results in a throughput gain of 30% over shortest path. Throughput increase compared to SPATH-CODE without listening is 10%, which is not significant. It is interesting to note that for this topology, CA-MPATH-CODE without opportunistic listening results in a better throughput than SPATH-CODE with opportunistic listening. However, increase over SPATH-CODE is only around 10%. Thus, for this topology, only the combined gain of both coding-aware routing and opportunistic listening is significant (40% over SPATH-CODE).

**Evaluation #3—Broadcast transmission bandwidth and throughput:** We compare the amount of broadcast transmission bandwidth required per unit throughput for: 1) MPATH; 2) SPATH-CODE with opportunistic listening; and 3) CA-MPATH-CODE with and without opportunistic listening. Fig. 10 shows that the number of broadcast transmissions normalized with respect to that of shortest-path routing decreases for SPATH-CODE up to a certain point and then remains constant. The initial decrease is due to the saving from increased coding opportunities (arising from increasing demands). We observe that MPATH requires around 15% more transmission bandwidth than shortest path, as it tries to increase throughput by routing along multiple paths, which might be longer but less interference-prone. SPATH-CODE requires an average bandwidth of only 0.83 per unit throughput, and the savings of around 17% compared to shortest path routing are purely due to coding. CA-MPATH-CODE without opportunistic listening requires a bandwidth of 0.94 even when it chooses potentially longer routes because of the reduced transmissions due to increased coding opportunities. CA-MPATH-CODE with opportunistic listening performs the best with 0.87, which is slightly higher than SPATH-CODE, but as observed, results in a higher throughput gain.

**Evaluation #4—Increase in throughput with fraction of coded traffic:** Fig. 11 shows that the fraction of coded traffic for SPATH-CODE and CA-MPATH-CODE both with and without opportunistic listening. As the number of flows increases, the fraction of coded traffic increases and then attains a maximum. Fig. 12 shows that throughput gain increases with fraction of coded traffic for SPATH-CODE, thus confirming the fact that observed gains are purely due to coding opportunities.

*In-Building Mesh Network:* Fig. 7 shows a 32-node topology derived from a mesh network testbed [10]. The average node degree was 6.8, and the maximum degree was 14. The effective link rates were derived from the delivery ratios, and the shortest path routing was based on the ETX metric.

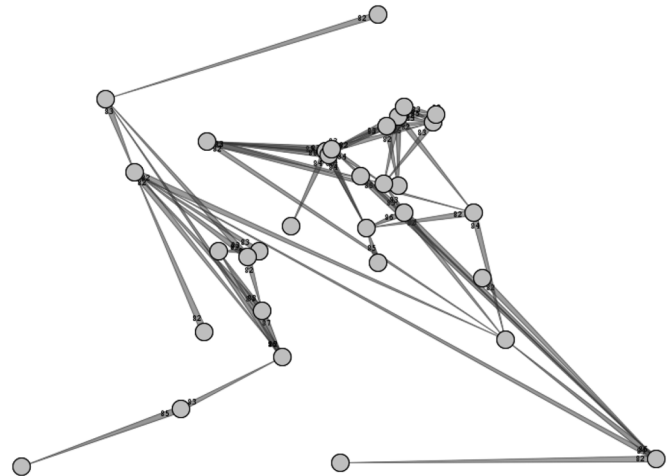


Fig. 6. Community wireless network.

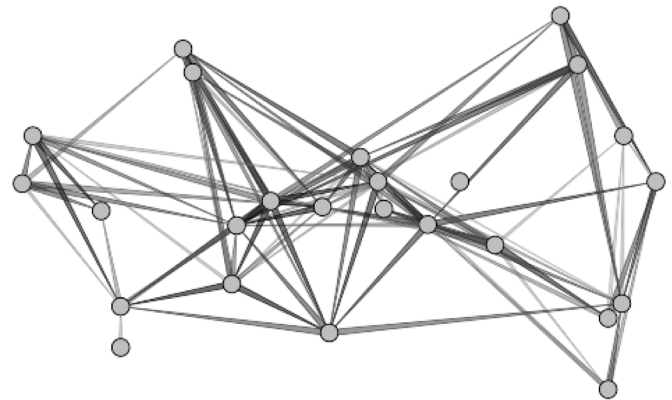


Fig. 7. In-building mesh network.

**Evaluation #5—Impact of network structure:** We computed the throughput gains for each of the schemes for increasing number of flows. Fig. 13 shows an interesting result: Throughput gains for SPATH-CODE and CA-MPATH-CODE are as high as 61% and 72% when opportunistic listening is allowed, while the corresponding gains without opportunistic listening are only 18% and 32%. We observe that in a network with a relatively high average degree of nodes, opportunistic listening facilitates increased coding opportunities and thus results in significant throughput improvements.

*Community Wireless Network:* We repeated the above evaluation on a 31-node topology derived from a campus-wide community wireless network (Fig. 6). The average node degree was 3.8, and the maximum degree was 8. Fig. 14 shows the throughput gains for SPATH-CODE and CA-MPATH-CODE both with and without opportunistic listening. We observe that the trend is similar to that observed in throughput gains for 15-node topology with the combined gains for opportunistic listening and coding-aware routing resulting in a throughput increase of 45%. We note that the increasing trend of the throughput numbers in Figs. 13 and 14 is due to the fact that the number of demands (300) did not increase the load on the network to a point where the throughput starts to decrease.

**Evaluation #6—Coding structures and throughput contributions:** We computed the average amount of coded traffic

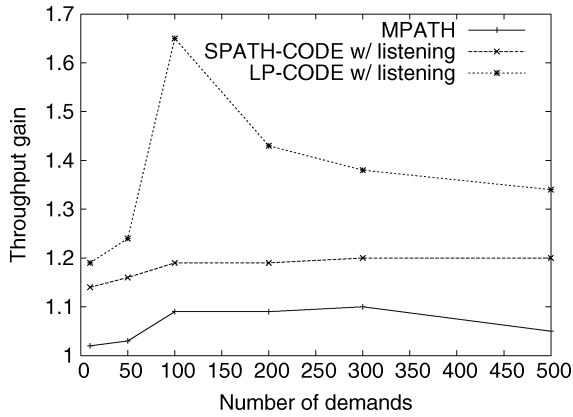


Fig. 8. Throughput gain (normalized w.r.t. SPATH).

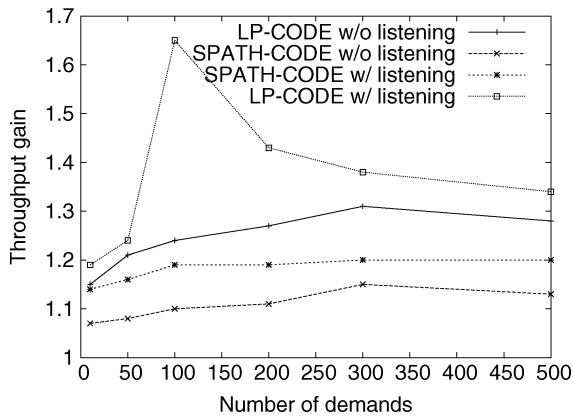


Fig. 9. Throughput gain (with and without opportunistic listening).

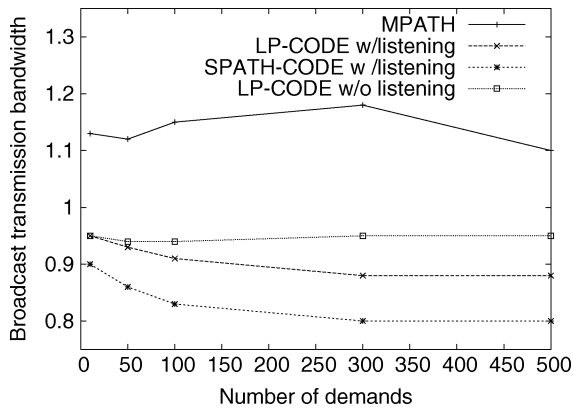


Fig. 10. Bandwidth/throughput normalized w.r.t. SPATH.

contributed by: 1) structures requiring opportunistic listening (S-OL); and 2) structures not requiring opportunistic listening (S-NOL) for each of the topologies mentioned. Fig. 15 shows that for the mesh network (Fig. 7), the majority of the throughput is contributed by coding structures requiring opportunistic listening, thus justifying the increased throughput gains for SPATH-CODE and CA-MPATH-CODE when opportunistic listening is allowed.

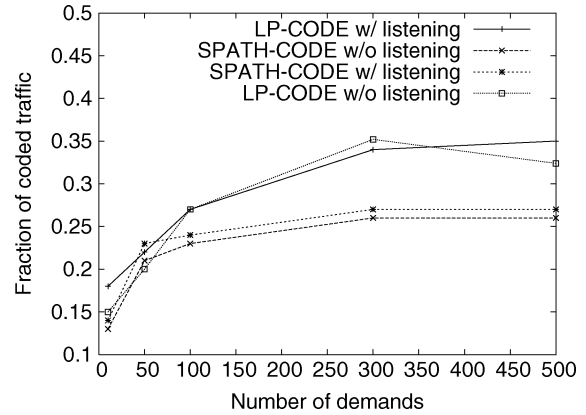


Fig. 11. Fraction of coded traffic versus number of demands.

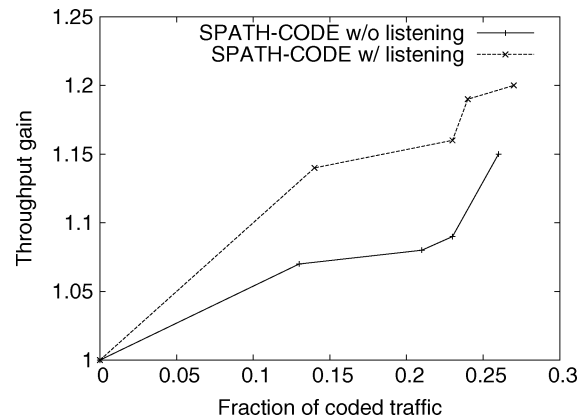


Fig. 12. Throughput gain with coded traffic.

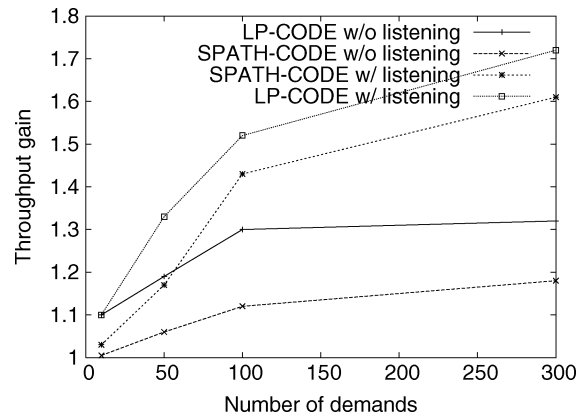


Fig. 13. Throughput gain (mesh network).

## VIII. CONCLUSION

In this paper, we have presented a theoretical framework for a detailed analytical evaluation of a practical network coding approach, such as COPE, for improving throughput in a multi-hop wireless network. Our formulations provide a systematic method to quantify the benefits of using network coding in the presence of multiple concurrent unicast sessions. We work with any wireless topology and any pattern of concurrent unicast traffic demands, and our results are valid both in the presence and absence of opportunistic listening mechanisms. In addition,

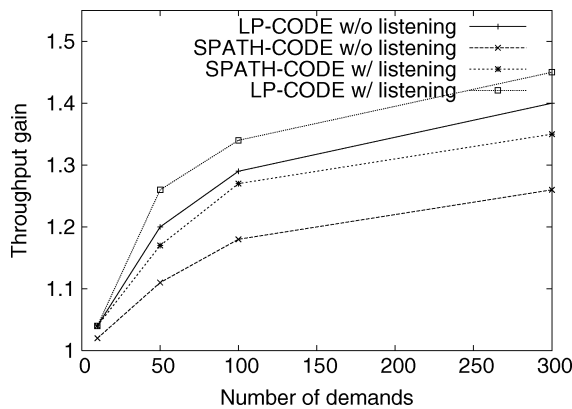


Fig. 14. Throughput gain (community wireless network).

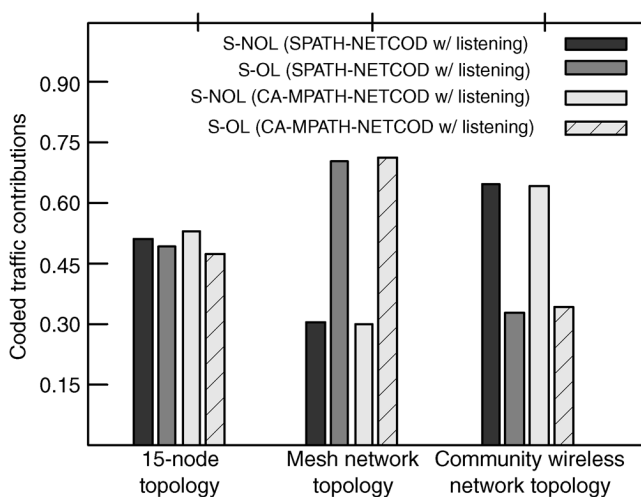


Fig. 15. Coded traffic associated with structures.

we introduced the notion of joint coding-aware and interference-aware routing for choosing routes that optimize the trade-offs between the conflicting effects of increased coding opportunities and increased wireless interference.

We evaluated our formulations on various topologies and found that coding-aware routing improved the throughput significantly, with observed gains reaching as high as 70% over traditional routing and 40% over coding-oblivious routing for some scenarios. However, we also observe that throughput improvements are dependent on the network structure, traffic pattern, and whether mechanisms like opportunistic listening are employed. We also showed that our model is extensible and the methodology used here is applicable to optimize many other objective functions such as minimizing total transmission bandwidth using network coding in wireless sensor networks.

Our theoretical formulations and evaluation results assume a “fair” MAC that does optimal scheduling so as to maximize throughput. The throughput improvements with network coding depend on the particular access mechanism employed by the MAC layer [12] and could be higher with an 802.11 like MAC, as reported in COPE [10]. These additional gains stem from inherent unfairness of 802.11 MAC. Adding a careful model for an 802.11 like MAC within our optimization framework is more challenging and will be the subject of our future work.

In conclusion, we hope that the framework presented in this paper can be a valuable tool to design protocols integrating network coding and routing selection techniques to increase end-to-end throughput in multihop wireless networks.

## REFERENCES

- [1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, “Network information flow,” *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.
- [2] P. A. Chou, Y. Wu, and K. Jain, “Practical network coding,” in *Proc. Allerton Conf. Commun.*, 2003.
- [3] D. S. J. De Couto, D. Aguayo, J. Bicket, and R. Morris, “A high-throughput path metric for multi-hop wireless routing,” in *Proc. ACM MobiCom*, 2003, pp. 134–146.
- [4] R. Dougherty, C. Freiling, and K. Zeger, “Insufficiency of linear coding in network information flow,” *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2745–2759, Aug. 2005.
- [5] R. Draves, J. Padhye, and B. Zill, “Routing in multi-radio, multi-hop wireless mesh networks,” in *Proc. ACM MobiCom*, 2004, pp. 114–128.
- [6] M. Gast, *802.11 Wireless Networks: The Definitive Guide*. Sebastopol, CA: O’Reilly, 2002.
- [7] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.
- [8] T. C. Ho, Y.-H. Chang, and K. J. Han, “On constructive network coding for multiple unicasts,” in *Proc. 44th Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2006.
- [9] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, “Impact of interference on multi-hop wireless network performance,” in *Proc. ACM MobiCom*, Sep. 2003, pp. 66–80.
- [10] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, “XORs in the air: Practical wireless network coding,” in *Proc. ACM SIGCOMM*, Sep. 2006, pp. 243–254.
- [11] R. Koetter and M. Medard, “An algebraic approach to network coding,” *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, Oct. 2003.
- [12] J. Le, J. C. S. Lui, and D. M. Chiu, “How many packets can we encode?—An analysis of practical wireless network coding,” in *Proc. IEEE INFOCOM*, Apr. 2008, pp. 371–375.
- [13] Z. Li and B. Li, “On increasing end-to-end throughput in wireless ad hoc networks,” in *Proc. QShine*, 2005, pp. 25–33.
- [14] Z. Li and B. Li, “Network coding: The case for multiple unicast sessions,” in *Proc. Allerton Conf. Commun.*, 2004.
- [15] S.-Y. R. Li, R. W. Yeung, and N. Cai, “Linear network coding,” *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.
- [16] D. S. Lun, M. Medard, and R. Koetter, “Efficient operation of wireless packet networks using network coding,” in *Proc. IWCT*, Jun. 2005.
- [17] D. S. Lun, N. Ratnakar, R. Koetter, M. Médard, E. Ahmed, and H. Lee, “Achieving minimum cost multicast: A decentralized approach based on network coding,” in *Proc. IEEE INFOCOM*, Mar. 2005, vol. 3, pp. 1607–1617.
- [18] M. Medard, M. Effros, T. Ho, and D. Karger, “On coding for non-multicast networks,” in *Proc. Allerton Conf. Commun.*, 2003.
- [19] A. Ramamoorthy, J. Shi, and R. D. Wesel, “On the capacity of network coding for random networks,” in *Proc. Allerton Conf. Commun.*, 2003.
- [20] S. Sengupta, S. Rayanchu, and S. Banerjee, “An analysis of wireless network coding for unicast sessions: The case for coding-aware routing,” in *Proc. IEEE INFOCOM*, May 2007, pp. 1028–1036.
- [21] D. Traskov, M. Heindlmaier, M. Medard, R. Koetter, and D. S. Lun, “Scheduling for network coded multicast: A conflict graph formulation,” in *Proc. IEEE GLOBECOM*, Dec. 2008, pp. 1–5.
- [22] D. Traskov, N. Ratnakar, D. S. Lun, R. Koetter, and M. Medard, “Network coding for multiple unicasts: An approach based on linear optimization,” in *Proc. IEEE ISIT*, Jul. 2006, pp. 1758–1762.
- [23] Y. Wu, S. Das, and R. Chandra, “Routing with a Markovian metric to promote local mixing,” in *Proc. IEEE INFOCOM*, May 2007, pp. 2381–2385.
- [24] K. Zeng, W. Lou, and H. Zhai, “On end-to-end throughput of opportunistic routing in multirate and multihop wireless networks,” in *Proc. IEEE INFOCOM*, Apr. 2008, pp. 816–824.
- [25] “ILOG CPLEX,” [Online]. Available: <http://www.ilog.com>
- [26] A. Eryilmaz and D. S. Lun, “Control for inter-session network coding,” in *Proc. ITA*, Jan./Feb. 2007.
- [27] J. K. Sundararajan, M. Medard, R. Koetter, and E. Erez, “A systematic approach to network coding problems using conflict graphs,” in *Proc. ITA*, Feb. 2006.

- [28] J. K. Sundararajan, M. Medard, M. Kim, A. Eryilmaz, D. Shah, and R. Koetter, "Network coding in a multicast switch," in *Proc. IEEE INFOCOM*, May 2007, pp. 1145–1153.

**Sudipta Sengupta** (SM'07) received the B.Tech. degree in computer science and engineering from the Indian Institute of Technology (IIT), Kanpur, India, and the M.S. and Ph.D. degrees in electrical engineering and computer science from Massachusetts Institute of Technology (MIT), Cambridge.

He is currently with Microsoft Research, Redmond, WA. Previously, he spent five years at Bell Laboratories, Lucent Technologies. He has published numerous research papers in some of the top conferences, journals, and technical magazines. He has taught advanced courses on networking at academic/research and industry conferences. He has authored patents in the area of computer networking. His research interests include algorithms/optimization, system architecture, protocols, and software for data networking, wireless networking, and network security.

Dr. Sengupta won the IEEE Communications Society Leonard G. Abraham Prize in 2008 for his work on oblivious routing of Internet traffic. At Bell Labs, he received the President's Teamwork Achievement Award for technology transfer of research into Lucent products. His work on P2P-based distribution of real-time layered video received the IEEE ICME 2009 Best Paper Award. He was also awarded the President of India Gold Medal at IIT-Kanpur for graduating at the top of his class across all disciplines.

**Shravan Rayanchu** received B.Tech. degree in computer science and engineering from the Indian Institute of Technology, Guwahati, India, and the M.S. degree in computer science from the University of Wisconsin, Madison. He is a Ph.D. candidate in the Computer Science Department of the University of Wisconsin–Madison.

His research focuses on improving the performance of wireless networks by designing smart control and data plane mechanisms.

Mr. Rayanchu is a recipient of the 2009–2011 Microsoft Research Ph.D. fellowship. He also won the Best Paper Award for work on hybrid scheduling at ACM MobiCom 2009.

**Suman Banerjee** (M'05) received the B.Tech. degree in computer science and engineering from the Indian Institute of Technology, Kanpur, India, and the M.S. and Ph.D. degrees in computer science from the University of Maryland, College Park.

He is an Associate Professor with the Department of Computer Sciences, University of Wisconsin-Madison, and heads the newly founded Wisconsin Wireless and NetworkinG Systems (WiNGS) Laboratory. His research focus is on various aspects of wireless and mobile networking systems.

Dr. Banerjee received the NSF CAREER Award in 2008.