# Challenges: Towards Truly Scalable Ad Hoc Networks

J.J. Garcia-Luna-Aceves[†]
jj@soe.ucsc.edu

Hamid R. Sadjadpour[‡]
hamid@soe.ucsc.edu

Zheng Wang[‡]
wzgold@soe.ucsc.edu

Department of Computer Engineering [†] and Electrical Engineering [‡]
University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA
[†] Palo Alto Research Center (PARC), 3333 Coyote Hill Road, Palo Alto, CA 94304, USA

## ABSTRACT

The protocols used in ad hoc networks today are based on the assumption that the best way to approach multiple access interference (MAI) is to avoid it. Unfortunately, as the seminal work by Gupta and Kumar has shown, this approach does not scale. Recently, Ahlswede, Ning, Li, and Yeung showed that network coding (NC) can attain the max-flow min-cut throughput for multicast applications in directed graphs with point-to-point links. Motivated by this result, many researchers have attempted to make ad hoc networks scale using NC. However, the work by Liu, Goeckel, and Towsley has shown that NC does not increase the order capacity of wireless ad hoc networks for multi-pair unicast applications. We demonstrate that protocol architectures that exploit multi-packet reception (MPR) *do* increase the order capacity of random wireless ad hoc networks by a factor $\Theta(\log n)$ under the protocol model. We also show that MPR provides a better capacity improvement for ad hoc networks than NC when the network experiences a single-source multicast and multi-pair unicasts. Based on these results, we introduce design problems for channel access and routing based on MPR, such that nodes communicate with one another on a many-to-many basis, rather than one-to-one as it is done today, in order to make ad hoc networks truly scalable.

## Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design Wireless Communication**]: [Computer-Communication Network]

## General Terms

Performance, Theory

## Keywords

Multipacket Reception, Network Coding, Ad Hoc Networks, Unicast Capacity, Multicast Capacity, Multihop Wireless Networks

## 1. INTRODUCTION

The communication protocols used today in ad hoc networks are based on a one-to-one communication paradigm in which a given receiver is able to decode at most one transmission correctly and transmitters and receivers orchestrate transmissions trying to offer at most one transmission around a receiver at any given time. The main objective of this one-to-one communication approach is the avoidance of multiple access interference (MAI). Unfortunately, the seminal work by Gupta and Kumar [1] demonstrated that the per source-destination throughput in a connected random wireless ad hoc network of $n$ nodes adhering to such a communication paradigm scales as $\Theta\left(\frac{1}{\sqrt{n\log(n)}}\right)$[1]

for multi-pair unicast applications. This result was obtained under the protocol model [1], in which a transmission carries a single packet, and a given transmission is successful at a receiver only if the transmitter is within the reception range of the receiver and no other node transmits within a distance equal to $(1 + \Delta)$ times the reception range, where $\Delta$ is a function of the physical layer. Intuitively, the sharp decrease in capacity experienced as the number of nodes increases in an ad hoc network using point-to-point communication can be explained in the protocol model by the fact that a single successful transmission occupies a circumference given by the reception radius of the receiver, and this area is a function of the minimum radius needed for the network to be connected. Hence, as nodes are added, a smaller percentage of nodes are free to become a successful transmitter. Clearly, without exploiting node mobility [2,3], the only two possible approaches to increase the order capacity of an ad hoc network consist of (a) increasing the amount of information a transmitting node relays in each transmission, or (b) enabling a receiver to decode multiple concurrent transmissions within its reception radius. Work has been carried out in both fronts.

Recently, Ahlswede et al. introduced the concept of network coding (NC) [4], which allows nodes to conduct processing and combining on received packets before forwarding them. They proved that the max-flow min-cut throughput can be achieved for single source multicast applications in a directed graph in which there are no restrictions on when a node can send and receive information. This result has motivated a large number of researchers to investigate how to increase the throughput capacity of ad hoc networks us-

---

[1]$\Theta$, $\Omega$ and $O$ are the standard order bounds, and $\log(\cdot)$ is the natural logarithm.

ing NC (e.g., [5]). However, Liu et al. [6] recently showed that NC cannot increase the throughput order of wireless ad hoc networks for multi-pair unicast applications when nodes are half-duplex. This result is related to the theorem given by Li and Li [7], who proved that NC has no throughput gain for unicast and broadcast applications, and can provide a throughput capacity that is at most twice that of an undirected graph with no NC.

On the other hand, Ghez et al. [8, 9] and Mergen and Tong [10] provided a framework for many-to-one communication. In this context, multiple nodes cooperate to transmit their packets simultaneously to a single node using multiuser detection (MUD), directional antennas (DA), or multiple input multiple output (MIMO) techniques (e.g., [11, 12]). The receiver node utilizes MUD and successive interference cancellation (SIC) to decode multiple packets. Toumpis and Goldsmith [13] have shown that the capacity regions for ad hoc networks are significantly increased when multiple access schemes are combined with spatial reuse (i.e., multiple simultaneous transmissions), multi-hop routing (i.e., packet relaying), and SIC.

The first contribution of this paper is to demonstrate that the per source-destination throughput $T(n)$ of a random wireless ad hoc network of three dimensions (or 3-D network) in which nodes utilize MPR is bounded by $\Theta(r(n))$ (upper and lower bounds) w.h.p.[2] when the protocol model is used, where $r(n)$ is the reception range of a receiver. This result is presented in Section 3, and is quite remarkable and yet intuitive! First, we note that, to ensure connectivity in random wireless ad hoc networks, $r(n) \geq \Theta\left(\sqrt[3]{\log n/n}\right)$. This results in an achievable bound of $\Theta\left(\sqrt[3]{\log n/n}\right)$ for a random wireless ad hoc network endowed with MPR in its nodes, which represents a gain in the order capacity of $\Theta(\log(n))$ compared to that attained with simple multihop routing [1, 14] (one-to-one communication) or NC [6]. Second, our result is in stark contrast to all existing results in ad hoc networks assuming point-to-point communications! It states that *increasing* the communication range $r(n)$ actually *increases* the capacity of an ad hoc network. Intuitively, the reason for this is that, given that all receivers are endowed with MPR, MAI around any receiver becomes useful information and no longer decreases the capacity.[3] Clearly, the restrictions in choosing the communication range among nodes are: (a) the need to maintain connectivity in the network, which provides a lower bound on $r(n)$; and (b) the energy spent per transmission, the transmitter complexity, and the decoding complexity of the nodes in the network, which provides a practical upper bound on $r(n)$.

The second contribution of this paper consists of showing that, in a wireless ad hoc network in which optimum routing is known to all sources and a combination of multi-pair unicasting and a single-source multicasting to a small group take place, MPR provides a gain in the order capacity of the network compared to NC. This result is addressed in Section 4. Intuitively, it can be explained by noticing that the percentage of multicast traffic becomes smaller compared to the multi-pair unicasts as the number of nodes increases, and then applying the results by Liu et al. [6] and our result from Section 3. We conjecture that the highest attainable

capacity gain for multiple source multicast can be achieved when MPR and NC are combined together. Computing the capacity of multiple source multicasts in networks that utilize MPR is tedious but can be shown using the techniques described in this paper. Computing the same capacity for the case of NC remains an open problem.

Given the above results and recent work related to the use of MPR in ad hoc wireless networks [15, 16], Section 5 addresses the use of MPR in practice, and introduces open problems in channel access and routing that take advantage of MPR in wireless ad hoc networks.

We hope that this paper motivates the design and analysis of protocol architectures for ad hoc networks that exploit rather than avoid MAI by means of MPR as a new challenge whose solution(s) can one day render truly scalable ad hoc networks.

## 2. PRELIMINARIES AND RELATED WORK

Considerable attention has been devoted to improving or analyzing the landmark results by Gupta and Kumar [1] on the scalability of wireless networks, and we only mention a very small fraction of these works due to space limitations. Gupta and Kumar [14] extended their own work to 3-D and K-D models. The gap between lower and upper bounds for these networks based on the physical model was closed [17] using percolation theory. It has been shown [18] that, if the physical layer assumption such as bandwidth expansion is changed, a throughput capacity of $O(n \log n)^{(\alpha-1)/2}$ and $\Omega(\frac{n^{(\alpha-1)/2}}{(\log n)^{(\alpha+1)/2}})$ can be attained for upper and lower bounds, respectively. Zhang and Hou closed the gap of such bounds applying percolation theory [19]. Moreover, it has been shown that mobility achieves non-zero capacity [2] for wireless ad hoc networks, and this work has been extended to one-dimensional (1-D) [20] and restricted mobility [21]. Jacquet and Rodolakis [22] and Shakkottai et al [23] have addressed the multicast capacity of wireless networks when nodes employ traditional multihop routing.

Our analysis focuses on the 3-D model of wireless ad hoc networks in which nodes are endowed with MPR capability. Our model is consistent with the analytical model in [14], and considers networks represented with an undirected graph (bidirectional links) such that two nodes $X_i$ and $X_j$ can communicate directly only if they are connected with an edge. This graph model has traditionally been used assuming a collision channel [1], which we also call the one-to-one communication assumption. That is, two nodes can communicate directly if they are within a distance $d(n)$, and the transmission from node $X_i$ to node $X_j$ is successful only if there is no other transmitter within distance $(1+\Delta)d(n)$ to node $X_j$, where $\Delta$ is a parameter that depends on the characteristics of the physical layer. This is called the protocol model [1], and inherently implies that the disks of different concurrent receivers with radius $d(n)$ are disjoint.

Applying the same protocol model to wireless networks with MPR capability means that nodes are able to receive successfully multiple packets concurrently, as long as the transmitters are within a radius of $r(n)$ from the receiver and all other transmitting nodes have a distance larger than $(1 + \Delta)r(n)$. The key difference is that MPR allows the receiver node to receive multiple packets from different nodes within its disk of radius $r(n)$ simultaneously. Note that $d(n)$ in point-to-point communication is a random variable

---

while $r(n)$ in MPR is a predefined value. To consider such networks, we use the graph models with MPR [8–10]. We assume that a node cannot transmit and receive at the same time and that it can transmit only one packet at a time, which is the norm in wireless ad hoc networks in practice.

Before proceeding with our discussion of capacity limits, we need to introduce a few results that we will use in our computations. First, Gupta and Kumar [14] showed that the connectivity among nodes in the 3-D model is guaranteed w.h.p. if and only if the transmit communication range $r(n)$ is lower bounded by

$$r(n) \geq \Theta \left( \sqrt[3]{\frac{\log n}{n}} \right). \qquad (1)$$

They also showed [14] that the distribution of nodes in random networks is uniform, so if there are $n$ nodes in a unit volume, then the density of nodes equals $n$. Hence, if $|V|$ denotes the volume of space region $V$, the expected number of the nodes, $E(N_V)$, in this volume is given by

$$E(N_V) = n|V|, \qquad (2)$$

Let $N_j$ be a random variable defining the number of nodes in $V_j$. Then, for the family of variables $N_j$, we have the following standard results known as the Chernoff bounds [24]:

- For any $\delta > 0$, $P[N_j > (1+\delta)n|V_j|] < \left( \frac{e^\delta}{(1+\delta)^{1+\delta}} \right)^{n|V_j|}$

- For any $0 < \delta < 1$, $P[N_j < (1-\delta)n|V_j|] < e^{-\frac{1}{2}n|V_j|\delta^2}$

Combining these two inequalities we have, for any $0 < \delta < 1$:

$$P[|N_j - n|V_j|| > \delta n|V_j|] < e^{-\theta n|V_j|} \qquad (3)$$

where $\theta = (1+\delta)\ln(1+\delta) - \delta$ in the case of the first bound, and $\theta = \frac{1}{2}\delta^2$ in the case of the second bound. Therefore, for any $\theta > 0$, there exist constants such that deviations from the mean by more than these constants occur with probability approaching zero as $n \to \infty$. It follows that, w.h.p., we can get a very sharp concentration on the number of nodes in a volume, so we can find the achievable lower bound w.h.p., provided that the upper bound is given. In the next section, we first derive the upper bound, and then use the Chernoff bound to prove the achievable lower bound w.h.p.

Lastly, we note that the capacity results that we present also depend on the transmission bandwidth $W$ of the network. However, given that we assume that $W$ is independent of $n$, the value of $W$ is simply a constant multiplier in capacity computations and does not change the order of capacity. Hence, we consider $W = 1$ for simplicity.

## 3. THROUGHPUT CAPACITY OF MULTI-PAIR UNICASTING

We now derive the capacity for the case of multi-pair unicasts in a 3-D random wireless ad hoc network in which nodes use MPR. For convenience, we refer to this case as the MPR scheme or the MPR protocol model. To simplify our analysis, we assume that $n$ nodes are randomly located inside a cube of a unit area[4]. Each node selects a destination randomly, and sends $T(n)$ bits/sec.

[4]In order to avoid edge effects, we could use a sphere as in [14] and the results of this paper would not change.

### 3.1 Upper Bound

A cut $\Gamma$ is a partition of a network into two connected components. The cut capacity is defined as the sum of bandwidth of all the edges crossing the cut. A min-cut is a cut whose capacity is the minimum value of all cuts. For wireless networks, we use a *sparsity cut* instead of min-cut to take into account the broadcast nature of wireless links [6].

In the 3-D case, the cut plane $\Gamma_p$ is defined as the area of the cut. The cut plane that we consider has zero width, such that no node lies on it. A sparsity cut for a random network is defined as a cut induced by the plane with the minimum area that separates the region into two subregions [6]. For the cubic deployment region illustrated in Fig. 1, the middle plane induces a sparsity cut $\Gamma_p$. Because nodes are uniformly deployed in a random network, such a sparsity cut captures the traffic bottleneck of a random network on average. This cut capacity constrains the information rate that the nodes from one side of the cut as a whole can deliver to the nodes at the other side. This is the maximum information (in bits per second) that can be transmitted across the cut from left to right (or from right to left).
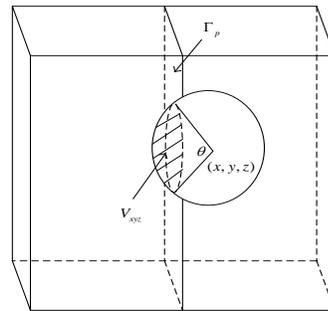


**Figure 1: Cubic unit area.**

The sparsity cut capacity is upper bounded by deriving the maximum number of simultaneous transmissions across the cut. In the work by Gupta and Kumar [14], spheres of radius $r(n)$ centered at each receiver are not necessarily disjoint, and the protocol model is still satisfied as long as the transmitter has the closest distance to the receiver node compared to all other transmitters in the network. In the MPR protocol model, the receiver node can receive packets simultaneously from all the nodes within a radius of $r(n)$ and all other transmitters should be outside of region of radius $r(n)$. More specifically, the MPR protocol model allows simultaneous transmissions by multiple nodes around a receiver as long as they are within a radius of $r(n)$ from the receiver and all other transmitting nodes have a distance larger than $(1+\Delta)r(n)$. $\Delta$ is a guard zone that is a function of the physical layer characteristics.

LEMMA 3.1. *The capacity of a sparsity cut $\Gamma$ for a unit region has an upper bound of $2c_1\Gamma_p r(n)n$, where $c_1 = \frac{1}{3(1+\frac{\Delta}{2})^2}$.*

PROOF. The cut capacity is upper bounded by the maximum number of simultaneous transmissions across the cut. In Fig. 1 we observe that all the nodes located in the shaded volume $V_{xyz}$ can send their packets to the receiver node located at $(x, y, z)$. These nodes lie in the left side of the cut $\Gamma_p$ within an area called $V_{xyz}$ and the assumption is that all these nodes send packets to the right side of the cut $\Gamma_p$.

For a node at location $(x, y, z)$, any node in the sphere of radius $r(n)$ can transmit information to this node simultaneously and the node can successfully decode those transmissions. To obtain an upper bound, we only need to consider edges that cross the cut. We first consider all possible nodes that can transmit to the receiver node in the $V_{xyz}$ region. We use Eq. (2) to estimate the average number of transmitters located in $V_{xyz}$ as $nV_{xyz}$. The number of nodes that are able to transmit at the same time from left to right is upper bounded as a function of $V_{xyz}$. The volume of $V_{xyz}$, which is a spherical cap, is given by

$$
\begin{aligned}
V_{xyz} &= \frac{1}{6}\pi r^3(n)\left(1 - \cos\frac{\theta}{2}\right)\left(3\sin^2\frac{\theta}{2} + \left(1 - \cos\frac{\theta}{2}\right)^2\right) \\
&= \frac{1}{3}\pi r^3(n)\left(1 - \cos\frac{\theta}{2}\right)^2\left(2 + \cos\frac{\theta}{2}\right).
\end{aligned} \quad (4)
$$

Hence, the total number of nodes that can send packets across the cut is

$$
\begin{aligned}
N_{max} &= \max_{0 \le \theta \le \pi}\left[\frac{\Gamma_p}{\pi(1+\frac{\Delta}{2})^2 r^2(n)\sin^2\frac{\theta}{2}}V_{xyz}n\right] \\
&= \max_{0 \le \theta \le \pi}\left[c_1\Gamma_p\frac{(1 - \cos\frac{\theta}{2})(2 + \cos\frac{\theta}{2})}{1 + \cos\frac{\theta}{2}}r(n)n\right]
\end{aligned} \quad (5)
$$

where $c_1 = \frac{1}{3(1+\frac{\Delta}{2})^2}$. This number is maximized with $\theta = \pi$. Therefore, the total number of nodes is upper bounded by $2c_1\Gamma_p r(n)n$. $\square$

COROLLARY 3.2. *For any unit-volume 3-D random network of arbitrary shape, if the minimum cut plane $\Gamma_p$ is not a function of $n$, then the sparsity cut capacity has an upper bound of $O(nr(n))$.*

PROOF. Regardless of the shape of the unit volume region, there exists a sparsity cut for each orientation of the cut plane. This sparsity cut capacity depends only on the minimum cut area $\Gamma_p$. If $\Gamma_p$ is not a function of $n$, then the capacity is always upper bounded by $O(nr(n))$. $\square$

THEOREM 3.3. *The per source-destination throughput of the MPR scheme in a 3-D random network is upper bounded by $O(r(n))$.*

PROOF. For a sparsity cut $\Gamma$ in the middle of the network, on average, there are $\Theta(n)$ pairs of source-destination nodes that need to cross $\Gamma$ in one direction w.h.p., i.e., $n_{\Gamma_{1,2}} = n_{\Gamma_{2,1}} = \Theta(n)$. The theorem then follows by combining this result with Corollary 3.2. $\square$

## 3.2 Lower Bound

In this section, we prove that, when $n$ nodes are distributed uniformly over a unit cubic volume, there are simultaneously at least $\frac{c_2\Gamma_p}{r^2(n)}$ circular regions, where $c_2 = \frac{1}{\pi(1+\Delta/2)^2}$, and each such region contains $\frac{2}{3}n\pi r^3(n)$ nodes w.h.p. This allows us to obtain an achievable lower bound by using the Chernoff bound, such that the distribution of the number of edges across the cut plane is sharply concentrated around its mean. Therefore, in a randomly chosen network, the actual number of edges crossing the sparsity cut plane is indeed $\Theta(nr(n))$ w.h.p.

THEOREM 3.4. *Each spherical region $V_j$ contains $\Omega(nr^3(n))$ nodes w.h.p. for all values of $j, 1 \le j \le \frac{c_2\Gamma_p}{r^2(n)}$.*

This theorem can be expressed as

$$
\lim_{n\to\infty} P\left[\bigcap_{j=1}^{c_2\Gamma_p/r^2(n)} |N_j - E(N_j)| < \delta E(N_j)\right] = 1, \quad (6)
$$

where $\delta$ is a positive small value arbitrarily close to zero.

PROOF. From the Chernoff bound and Equation (3), for any given $0 < \delta < 1$, we can find $\theta > 0$ such that

$$
P\left[|N_j - E(N_j)| > \delta E(N_j)\right] < e^{-\theta E(N_j)} = e^{-\theta n|V_j|} \quad (7)
$$

Thus, we can conclude that the probability that the value of the random variable $N_j$ deviates by an arbitrarily small constant value from the mean tends to zero as $n \to \infty$. Thus, when all the events $\bigcap_{j=1}^{c_2\Gamma_p/r^2(n)} |N_j - E(N_j)| < \delta E(N_j)$ occur simultaneously, then all $N_j$'s converge uniformly to their expected values. Utilizing the union bound, we obtain

$$
\begin{aligned}
&P\left(\bigcap_{j=1}^{c_2\Gamma_p/r^2(n)} |N_j - E(N_j)| < \delta E(N_j)\right) \\
&= 1 - P\left[\bigcup_{j=1}^{c_2\Gamma_p/r^2(n)} |N_j - E(N_j)| < \delta E(N_j)\right] \\
&\ge \max\left\{1 - \sum_{j=1}^{c_2\Gamma_p/r^2(n)} P\left[|N_j - E(N_j)| < \delta E(N_j)\right], 0\right\} \\
&> \max\left\{1 - \frac{c_2\Gamma_p}{r^2(n)}e^{-\theta E(N_j)}, 0\right\}.
\end{aligned} \quad (8)
$$

Because $E(N_j) = \frac{2\pi}{3}nr^3(n)$, the final result is

$$
\begin{aligned}
&\lim_{n\to\infty} P\left[\bigcap_{j=1}^{c_2\Gamma_p/r^2(n)} |N_j - E(N_j)| < \delta E(N_j)\right] \\
&\ge \max\left\{1 - \frac{c_2\Gamma_p}{r^2(n)}e^{-\frac{2\pi\theta}{3}nr^3(n)}, \ 0\right\}
\end{aligned} \quad (9)
$$

To guarantee connectivity, $r(n) > \sqrt[3]{\frac{\log n}{n}}$ [14]. Therefore, as $n \to \infty$, we have $\frac{e^{-\frac{2\pi\theta nr^3(n)}{3}}}{r^2(n)} \to 0$. $\square$

This theorem demonstrates that, w.h.p., we can achieve the lower bound.

COROLLARY 3.5. *The per source-destination throughput of MPR scheme for a 3-D random network has an achievable lower bound of $\Omega(r(n))$ w.h.p.*

PROOF. Theorem 3.4 proves that there are $\frac{\Gamma_p}{\pi(1+\Delta/2)^2 r^2(n)}$ different circles of radius $r(n)$, each of them having $\Theta(nr^3(n))$ nodes w.h.p. Therefore, the per source-destination throughput is the multiplication of these two values divided by the total number of nodes, which proves the corollary. $\square$

## 3.3 Selecting The Transmission Radius

From Theorem 3.3 and Corollary 3.5 for the upper and lower bounds of throughput capacity, respectively, the convergence of throughput capacity can be derived as $\Theta(r(n))$. Clearly, the selection of $r(n)$ is the key factor for the throughput capacity problem in static wireless ad hoc networks.

From Eq. (1), To guarantee connectivity of a 3-D network, the lower bound of $r(n)$ is $\Omega(\sqrt[3]{\log n/n})$. However, the upper bound of $r(n)$ can be any large value up to $O(1)$. Ultimately, the upper bound of $r(n)$ depends on the decoding complexity of the nodes in the network and the number of nodes that can be transmitted simultaneously. In practical wireless ad hoc networks, if the network is very dense, then the important value for $r(n)$ is its lower bound $\Omega(\sqrt[3]{\log n/n})$. That value is directly related to the connectivity of the network, which is given by Eq. (1). Constructively, therefore, the following theorem follows for the 3-D case.

THEOREM 3.6. *The per source-destination throughput $T(n)$ of the MPR scheme for a 3-D random network is given by $\Theta(r(n))$ as the lower and upper bounds w.h.p.*

# 4. THROUGHPUT CAPACITY WITH SINGLE-SOURCE MULTICASTING AND MULTI-PAIR UNICASTING

To simplify our presentation, we consider a 2-D network in this section. We first compute the capacity of a network in which a single-source multicast to a small group takes place, together with all other nodes participating in unicast communication. We use the term *hybrid routing* to denote such a combination of unicast and multicast routing in a network. Hence, in a network with hybrid routing, each node either participates in the single-source multicast or one of the unicasts in the network.
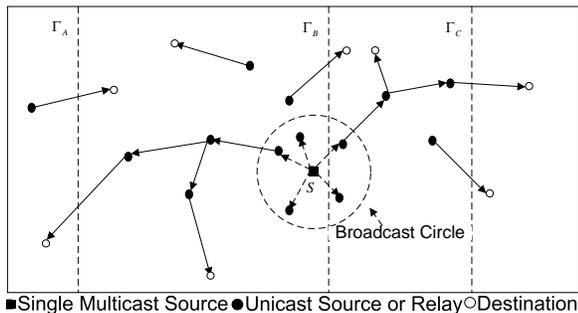


Figure 2: Two-step model in multicast application

In our proof, we take advantage of the fact that, because we address ad hoc networks with broadcast links, a single-source multicast can be viewed as a two-step forwarding process, which is illustrated in Fig. 2 and consists of the multicast source broadcasting a packet to the first-hop relays, followed by the throughput-optimum routing of the packet replicas from those first-hop relays to all the multicast destinations. For convenience we refer to throughput-optimum routing and throughput-optimum multicasting simply as optimum routing and optimum multicasting, respectively.

Given that we have computed the order capacity for multi-pair unicasts in Section 3, the realization that optimum single-source multicasting is a two-phase forwarding process

as stated above simplifies the problem of computing the multicast capacity for hybrid routing into a simpler and more manageable problem.

In the following theorems, $f(n)$ denotes the number of destinations in the single-source multicast taking place in a network with hybrid routing. Furthermore, it is assumed that the source of the multicast group is not at the edges of the network in order to avoid edge effects.

THEOREM 4.1. *In an ad hoc network with hybrid routing in which MPR is used at each node, the throughput capacity of the network is of order $\Theta(r(n))$ when $n \to \infty$, provided that $f(n)$ is such that $\lim_{n \to \infty} \frac{f(n)}{nr(n)} = \frac{f(n)}{\sqrt{n \log(n)}} \to 0$.*

PROOF. We define $\Gamma_A$, $\Gamma_B$, and $\Gamma_C$ to be three different cuts, as shown in Fig. 2. $\Gamma_A$ and $\Gamma_C$ are outside the transmission range of multicast source $S$, and $\Gamma_B$ is inside this transmission range of the multicast source.

For the case that the cut crosses the transmission range of the multicast source (cut $\Gamma_B$), we have two components to consider regarding the contributions to the simultaneous transmissions across the cut from one side (left) to another (right). One component consists of those nodes that do not participate in the multicast group and who transmit as part of the multi-pair unicasts. The other component consists of the multicast source and first-hop relays of the multicast group.

The first component consists of the multi-pair unicasts that we computed earlier. The only difference here is the fact that the circle around the multicast source should be excluded for computation of unicast contributions and it is given by (see Fig. 3)

$$\frac{\Gamma_B - 2r(n)\sin(\frac{\beta}{2})}{(2+\Delta)r(n)} \frac{1}{2}\pi n r^2(n). \tag{10}$$

It is clear that the asymptotic value of this equation is equal to $\Theta(nr(n))$ as $n \to \infty$.

The contribution of the multicast source and the first-hop relays to the capacity of the cut is based on the number of destinations in one side of the cut. This definition captures the fact that, even if the same packet is delivered to many destinations, all these deliveries are counted for capacity. If we assume that the multicast source is not at the edges of the network, then the number of destinations in each side of the cut is proportional to $f(n)$. Therefore, by adding these two values and dividing the result by $n$, the per source-destination capacity corresponding to this cut can be derived as $\Theta(r(n)) + \Theta(\frac{f(n)}{n})$. It is easy to see that this value is proportional to $\Theta(r(n))$ as $n \to \infty$ if $\frac{f(n)}{nr(n)} \to 0$.

Note that multicasting consists of a single broadcasting from the source to first-hop relays, followed by multi-pair unicasts, with the only difference being that each one of the first-hop relays may send data to more than one destination. Because $f(n)$ does not change the order capacity of these cuts as $n \to \infty$, it is clear that the capacity of these cuts is similar to the capacity for multi-pair unicasts, which we have shown to be proportional to $\Theta(r(n))$.

Because the locations of multicast destinations are randomly distributed, some of these destinations do not contribute to the capacity of cuts other than $\Gamma_B$, such as $\Gamma_A$ and $\Gamma_C$. However, we showed that, even when all the destinations are included in a cut, which is the case of $\Gamma_B$, the order capacity of this cut is still $\Theta(r(n))$. Hence, the per

source-destination capacities corresponding to $\Gamma_A$ and $\Gamma_C$, or any cut that does not cross the transmission range of the multicast source, are less than that of $\Gamma_B$ and of the same order of $\Theta(r(n))$, and can be computed from Theorem 3.6.

From Corollary 3.2, Theorem 3.3, and Theorem 3.4, we conclude that the achievable tight bound is $\Theta(r(n))$, which is the same as the capacity of Theorem 3.6. □
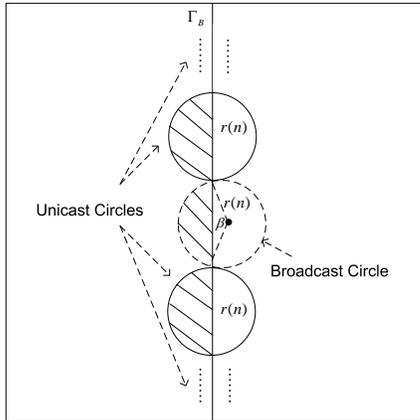


**Figure 3: The cut that crosses the broadcast circle**

THEOREM 4.2. *In an ad hoc network with hybrid routing in which NC is used, if $f(n)$ satisfies $\lim_{n\to\infty} \frac{f(n)\times\sqrt{log(n)}}{\sqrt{n}} = 0$, then the capacity of the network is similar to its capacity with only multi-pair unicasts.*

PROOF. In an ad hoc network with hybrid routing and in which NC is used, the total throughput that corresponds to the multicast operation is at most of order $\Theta(f(n))$ which is much smaller than the total number of nodes in the network. However, the total capacity of multi-pair unicasts is equal to $\frac{\sqrt{n}}{\sqrt{log(n)}}$ when NC is used [6]. It is obvious to see that the majority of capacity in this network is related to multi-pair unicasts similar to the results in Theorem 4.1. Because multicast operation is equivalent of a single broadcast and multi-pair unicasts as we showed earlier, the hybrid routing does not change the order of throughput in a network utilizing NC, as long as $\lim_{n\to\infty} \frac{f(n)\times\sqrt{log(n)}}{\sqrt{n}} = 0$. □

Note that the condition in Theorem 4.2 is stronger than the condition in Theorem 4.1. The main reason to choose a stronger condition in Theorem 4.2 is the fact that the capacity of multi-pair unicast is lower in NC than in MPR. Also note that we have implicitly assumed that $\lim_{n\to\infty} n - f(n) = n$, which is equivalent of saying that the number of destinations in multicast is much smaller than the total number of nodes in the network, which is an acceptable assumption. With the above theorems, we can state the following theorem.

THEOREM 4.3. *In a wireless ad hoc network with hybrid routing, using MPR renders a throughput order gain compared to using NC.*

PROOF. By modeling single-source multicast as a broadcast to first-hop relays followed by multi-pair unicasts, it

follows that the capacity order of a wireless ad hoc network using NC is comparable to that obtained in [6], which does not change the order capacity of wireless ad hoc network compared to the original results in [1]. Combining theorems 4.1, 4.2, and 3.6 implies that using MPR does increase the order capacity of wireless ad hoc networks by a factor of $log(n)$ compared to using NC. □

From the above, it appears that using MPR to make ad hoc networks scale for hybrid unicast and multicast applications is a better approach than NC. Figure 4(a) shows a network with two sources and two destinations. It can be seen that, with simple routing, one bit from each source can be transmitted to both destinations in four time slots. However, using NC or MPR requires only three time slots to deliver the same number of bits. On the other hand, by combining NC and MPR, two bits can be transmitted in just two time slots. Based on this simple example, we claim the following conjecture.

CONJECTURE 4.4. *For multiple source multicast applications in wireless ad hoc networks, the highest capacity gain can be achieved when MPR and NC are combined together.*

The heuristic justification of this conjecture derives from the fact that, when all the nodes in the network are endowed with MPR capability, all transmitting nodes around a receiver can transmit simultaneously. On the other hand, by allowing NC in the network, a node can transmit simultaneously all the bits it receives from its neighbors to their respected destinations or relays. Therefore, the dissemination of information is maximized during transmission using NC and during reception using MPR. Accordingly, it is natural to believe that combining the two techniques will increase and perhaps maximize the capacity of wireless ad hoc networks.

# 5. MAKING AD HOC NETWORKS SCALE

The results presented in the previous sections demonstrate that MPR can have a tremendous impact on the performance of future ad hoc networks, because it enables the protocol architectures of such networks to be based on many-to-many communication,[5] which takes advantage of the broadcast nature of wireless links and the peer-to-peer nature of nodal interaction. However, turning these theoretical results into practice represents a big challenge. In practice, receivers can decode only a finite number of concurrent transmissions, rather than all the transmission that occur within their reception range. Therefore, trade-offs are needed between the added efficiency attained by means of concurrent transmissions, and the added cost incurred by the complexity of the receivers that must process such transmissions.

Furthermore, the communication protocols used to date in ad hoc networks have been designed to avoid MAI, and are derivatives of protocols and architectures originally designed for wired networks based on point-to-point links. For example, today's popular IEEE 802.11 DCF can be viewed as attempting to emulate "Ethernets in the sky" in that at most one transmission is allowed to reach a receiver, and

___
[5]We use the term "many-to-many communication" [15] to refer to the orchestration among transmitters and receivers in order to exploit the MPR capabilities offered at the physical layer.
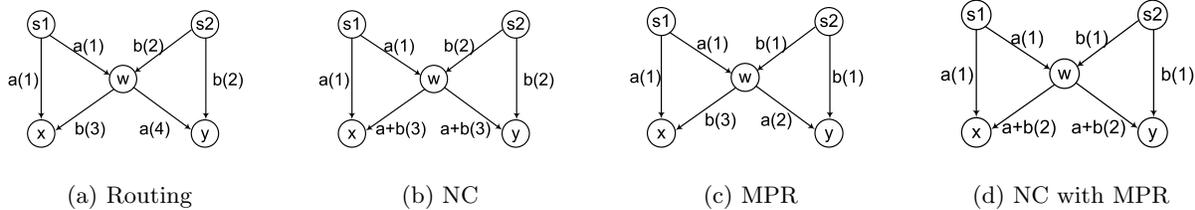
(a) Routing      (b) NC      (c) MPR      (d) NC with MPR

**Figure 4: The difference between Network Coding (NC) and Multipacket Reception (MPR)**

senders are forced to back off in the presence of MAI. Similarly, the IETF MANET routing protocols are based on the assumption that packets are to be forwarded along single paths, and they work independently of the channel access method, even though it is *not* true that routing in MANETs occurs over a pre-existing network topology and the transmission over one link does not impact the transmissions over other links, as it can be done in a wired network. Therefore, for MPR to really help ad hoc networks scale (i.e., an increase in capacity over today's approaches in the order of the degree of nodes in the network), the protocols used in such networks have to be redesigned from the ground up to *embrace*, not combat, MAI. While the entire protocol stack of an ad hoc network is impacted by a shift from one-to-one to many-to-many communication, we address only the link and network layers, because they are arguably the most affected by the change of communication paradigm.

## 5.1 Channel Access

For MPR to work in practice, the transmissions that must be decoded at a receiver need to be sent synchronously. Furthermore, because each receiver has finite processing power, the number of concurrent transmissions allowed around it cannot exceed a certain number that may be smaller than the number of neighbors near the receiver if the network is densely connected. Hence, taking advantage of MPR for channel access benefits from dynamic approaches to channel division (in time, frequency, and space), and requires the scheduling of concurrent transmissions around receivers based on their characteristics and the channel state at the receivers.

The receiver-oriented orchestration of transmissions desired for MPR can be accomplished using contention-based or scheduling approaches. However, we advocate a design consisting of establishing transmission schedules by means of distributed elections of the information flows that should be offered to the channel during the time slots of the schedule, together with opportunistic reservations of time slots after they are acquired through election. The rationale for election-based scheduling is that it can be carried out in a distributed manner and can be far more efficient than contention-based schemes. The motivation for opportunistic reservations is the need to support such traffic as voice, which is jitter and delay intolerant.

Clearly, simply stating the desirable features for channel access does not make a protocol! The design of MAC protocols that support many-to-many communication is a challenge that must be addressed in order to fully take advantage of MPR. Among the many problems that must be solved as part of this challenge, we can list the following: (a)

The coordination among senders and receivers under varying degrees of mobility, such that the receivers can provide feedback to senders on channel state information (CSI). (b) The efficient use of dynamic channel division (i.e., time division, frequency division, and space division) to divide the MAI around receivers, such that the probability of successful decoding by any one receiver is increased. (c) The incorporation of network-level information in the decisions made for elections and reservations (i.e., the integration of routing and scheduling). (d) The interaction between elections and reservations. (e) The integration of NC and MPR in the context of scheduling based on elections and reservations. While some work has been reported on the use of NC for channel access scheduling [25] the application of MPR and the combination of MPR and NC in scheduling deserve close attention.

## 5.2 Routing

Because the routing protocols designed for MANETs to date are based on avoiding MAI, they tend to maintain single paths to destinations, data packets are disseminated using single-copy forwarding, and protocol signaling attempts to minimize the number of nodes that forward information and MAI. Examples of techniques used to reduce MAI in routing and forwarding include introducing jitter in the transmission of periodic updates, using multiple node-disjoint paths to reach destinations, using multicast trees, and using multi-point relays to disseminate link states.

In contrast to the above, routing and forwarding in the context of many-to-many communication call for the exploitation of concurrency at the link level and redundancy at the network level, because the MAI caused by data and control packets can be managed and exploited. Hence, a route to a destination should be a "multipath" consisting of multiple paths that need not be edge or node disjoint, multi-copy forwarding [3] can be used to disseminate data over a multipath, periodic updates should be sent so that multiple concurrent updates reach the neighbors of transmitters at the same time, multicasting can be attained over "concurrency" meshes in which all multicast transmissions are useful information, and "feasible concurrency relays" can replace today's multi-point relays to disseminate control signaling in a way that a relay is feasible if it can transmit concurrently with other nodes to intended neighbors.

We believe that embedding the signaling needed to establish multipath routing with the signaling needed to establish transmission schedules is the main routing challenge for many-to-many communication. Important problems associated with this challenge are: (a) How should nodes elect and reserve time slots for the signaling required for unicasting

and multicasting that exploit concurrency and redundancy? (b) How should feasible concurrency relays be selected to maximize the reliability of signaling while making efficient use of link-level concurrency? (c) How should the "width" of the multipaths (number of neighbors each node uses as next hops to a destination) be controlled depending on demand? (d) How should reliability be increased or average delay or jitter be decreased by means of multi-copy forwarding over multipaths? and (e) How should MPR and NC be combined in the context of scheduled multipaths for the dissemination of data and control information? Note that the combination of packets needed for NC makes most sense over scheduled multipaths, because data packets must be forwarded along them. Hence, establishing the state needed to take advantage of NC can be very efficient if done as part of the signaling used to establish scheduled multipaths.

## 6. CONCLUSION

We have shown that exploiting MPR techniques in ad hoc networks can render a throughput capacity of order $\Theta\left(r(n)\right)$. The key significance of this result is that, with MPR, the ability of ad hoc networks to scale is no longer limited by MAI, but by the complexity of transmitters and receivers. Using recent results by Liu, Goeckel, and Towsley [6], we have also shown that MPR is a more attractive alternative than NC, insofar as making ad hoc networks scale with the number of nodes, and we have offered the conjecture that the combination of MPR and NC constitutes the best approach towards making ad hoc networks scale.

We have also argued that, in order to benefit from the scaling properties of MPR, the protocols of an ad hoc network must be based on many-to-many communication, and summarized key problems for the design of such protocols. We believe this to be an important challenge for the research community whose solution can render ad hoc networks that scale far better than today's networks.

## 7. ACKNOWLEDGMENTS

## 8. REFERENCES

[1] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[2] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks," in *Proc. of IEEE INFOCOM 2001*, Anchorage, Alaska, USA., April 22-26 2001.

[3] R. M. de Moraes, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves, "Throughput-delay analysis of mobile ad-hoc networks with a multi-copy relaying strategy," in *Proc. of IEEE SECON 2004*, Santa Clara, CA, USA., October 4-7 2004.

[4] R. Ahlswede, C. Ning, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[5] S. katti, D. Katabi, W. Hu, H. Rahul, and M. Medard, "The importance of being opportunistic: Practical network coding for wireless environments," in *Proc. of Allerton 2005*, Allerton House, UIUC, Illinois, USA., Sep. 2005.

[6] J. Liu, D. Goeckel, and D. Towsley, "Bounds on the gain of network coding and broadcasting in wireless networks," in *Proc. of IEEE INFOCOM 2007*, Anchorage, Alaska, USA., May 6-12 2007.

[7] Z. Li and B. Li, "Network coding in undirected networks," in *Proc. of CISS 2004*, Princeton, NJ, USA., March 17-19 2004.

[8] S. Ghez, S. Verdu, and S. C. Schwartz, "Stability properties of slotted aloha with multipacket reception capability," *IEEE Transaction on Automatic Control*, vol. 33, no. 7, pp. 640–649, 1988.

[9] ——, "Optimal decentralized control in the random access multipacket channel," *IEEE Transaction on Automatic Control*, vol. 34, no. 11, pp. 1153–1163, 1989.

[10] G. Mergen and L. Tong, "Stability and capacity of regular wireless networks," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1938–1953, 2005.

[11] P. Christina and D. S. Sergio, "On the maximum stable throughput problem in random networks with directional antennas," in *Proc. of ACM MOBIHOC 2003*, Annapolis, Maryland, USA., June 1-3 2003.

[12] Y. Su, P. Ying, and K. Sivakumar, "On the capacity improvement of ad hoc wireless networks using directional antennas," in *Proc. of ACM MOBIHOC 2003*, Annapolis, Maryland, USA., June 1-3 2003.

[13] S. Toumpis and A. J. Goldsmith, "Capacity regions for wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 736–748, 2003.

[14] P. Gupta and P. R. Kumar, "Internets in the sky: The capacity of three dimensional wireless networks," *Communications in Information and Systems*, vol. 1, no. 1, pp. 39–49, 2001.

[15] R. M. de Moraes, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves, "Many-to-many communication: A new approach for collaboration in manets," in *Proc. of IEEE INFOCOM 2007*, Anchorage , Alaska , USA., May 6-12 2007.

[16] A. Ozgur, O. Leveque, and D. Tse, "Hierarchical cooperation achieves linear capacity scaling in ad hoc networks," in *Proc. of IEEE INFOCOM 2007*, Anchorage , Alaska , USA., May 6-12 2007.

[17] M. Franceschetti, O. Dousse, D. N. Tse, and P. Thiran, "Closing the gap in the capacity of wireless networks via percolation theory," *IEEE Transactions on Information Theory*, vol. 53, no. 3, pp. 1009–1018, 2007.

[18] R. Negi and A. Rajeswaran, "Capacity of power constrained ad-hoc networks," in *Proc. of IEEE INFOCOM 2006*, Barcelona, Spain, April 23-29 2006.

[19] H. Zhang and J. Hou, "Capacity of wireless ad-hoc networks under ultra wide band with power constraint," in *Proc. of IEEE INFOCOM 2005*, Miami, FL, USA., March 13-17 2005.

[20] R. S. Diggavi, M. Grossglauser, and D. Tse, "Even one-dimentional mobility increases ad hoc wireless capacity," in *Proc. of IEEE ISIT 2001*, Washington, D.C. USA., June 24-29 2002.

[21] R. M. de Moraes, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves, "Mobility-capacity-delay trade-off in wireless ad hoc networks," *Ad Hoc Networks*, vol. 4, no. 5, 2006.

[22] P. Jacquet and G. Rodolakis, "Multicast scaling properties in massively dense ad hoc networks," in *IEEE ICPADS 2005*, Fukuoka, Japan, July 20-22 2005.

[23] S. Shakkottai, X. Liu, and R. Srikant, "The multicast capacity of wireless ad-hoc networks," in *Proc. of ACM MOBIHOC 2007*, Montreal, Canada, September 9-14 2007.

[24] R. Motwani and P. Raghavan, *Randomized Algorithms*. Cambridge University Press, 1995.

[25] Y. E. Sagduyu and A. Ephremides, "Joint scheduling and wireless network coding," in *Proc. of NetCod 2005*, Riva del Garda, Italy, Apr. 7 2005.