

THE CAPACITY OF MULTIHOP RELAY WIRELESS NETWORKS

Mingyue Ji[†], Zheng Wang[†], Hamid R. Sadjadpour[†], and J.J. Garcia-Luna-Aceves[‡]

Department of Electrical Engineering[†] and Computer 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
email: {davidjmy, wzgold, hamid, jj}@soe.ucsc.edu

ABSTRACT

The scaling laws of wireless networks are studied when there are $s(n)$ source-destination pairs in the network and the rest of $n - s(n)$ nodes are relays. The transmission strategy is based on multihop communications which is more practical in wireless ad hoc networks. We call this network *Multihop Relay Wireless Network (MRWN)*. It is shown that as the number of sources in the network increases (or equivalently the number of relays decreases), the capacity of such networks is either dominated by the number of source-destination pairs or by the number of relays. Both upper and lower bounds of throughput capacity are derived.

1. INTRODUCTION

The throughput capacity of wireless ad hoc networks is an important topic that has been studied extensively in the literature. Most of these studies [1–5] are based on the assumption that all the nodes in the network are either source or destination. However, there is another important group of networks where some nodes are neither source or destination. These nodes are considered relays and their task is to facilitate the transportation of information in the network. These networks are called relay networks. The capacity of relay networks has been studied in [6, 7]. The main assumption in these papers is that there are constant number of source-destination pairs which is not a function of the total number of nodes in the network. A more realistic assumption is to assume the number of source-destination and relays in the network is a function of the total number of nodes in the network. Further, the computation of throughput capacity in [6] is based on the assumption of two hop communications where the information is transmitted directly from sources to the relays and in the second hop, this information is transmitted to destinations. In a more realistic network, the information is transmitted through multiple hops from sources to destinations.

In this paper, we consider relay networks where the number of source-destination pairs, $s(n)$, is a function of the total number of the nodes in the network, n . The transportation of information from sources to destinations occurs through multiple hops depending on the distance between each source-destination pairs. We define this network as *Multihop Relay Wireless Network (MRWN)*. Since most wireless ad hoc networks such as military networks utilize multihop communications, this study is valuable in understanding the throughput capacity behaviour of these networks.

There are few works focusing on the network model similar to this paper. In [8], the authors assume data gathering

traffic where multiple sources send their independent data to a single destination while the rest of the nodes in the network perform as relays. The throughput capacity is derived by utilizing a multihop Manhattan routing protocol. In [9], the clustered network is considered where the density of the network is not uniform and varies. There are also other nodes distributed outside the clusters inside the network. All the information is generated by the clustered nodes and nodes outside the clusters perform as relays. In this work, the authors analyze the effects of the clustered nodes and the size of the network on the throughput capacity. In our work, all sources, destinations, and relays are distributed uniformly inside the network.

The rest of the paper is organized as follows. Section 2 presents assumptions and definitions used for our analysis. Section 3 provides the main result of this paper. We compute the upper bound in Section 4. The achievable throughput for *MRWN* is described in Section 5. Section 6 discusses the implications of the results. The paper is concluded in Section 7.

2. NETWORK MODEL

We consider a network with nodes uniformly distributed in a dense network, where the area of the network is a constant unit square. The total number of nodes in the network is n . There are $s(n)$ unicast traffic in the network and the other $n - s(n)$ nodes perform as relays. The transmission technique is point-to-point multihop routing scheme with no cooperation among nodes [4]. The transmission range, $r(n)$, is assumed to be the same for all the nodes. A successful communication between two nodes is modeled according to the protocol model [1], which is defined below.

Definition 2.1 *Protocol Model:* Node i at location X_i can successfully transmit to node $i(R)$ at location $X_{i(R)}$ with rate W bits/second if $|X_i - X_{i(R)}| \leq r(n)$ and for every node k located at X_k , $k \neq i$ that transmits at the same time, $|X_k - X_{i(R)}| \geq (1 + \Delta)r(n)$. $r(n)$ is the common transmission range in the network and Δ is related to the guard zone around the receiver.

This network model is shown in Fig 1.

The feasible throughput and the order of throughput capacity are defined as follows.

Definition 2.2 *Feasible Throughput:*

A throughput of $\lambda_i(n)$ bits per second is said to be feasible for the i^{th} source-destination pair if there is a common transmission range $r(n)$, and a scheme to schedule transmissions

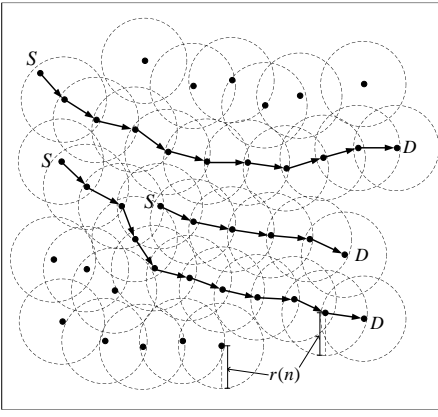


Figure 1: S and D denote source and destination respectively and other nodes are relays. The sources send packets to destinations in a multi-hop fashion. $r(n)$ is the common transmission range.

and there are routes between sources and destinations, such that source i can transmit to its destination at such rate successfully.

Definition 2.3 *Order of Throughput Capacity:* The total throughput capacity is said to be of order $\Theta(f(n))$ bits per second if there exists a constant c and c' such that

$$\lim_{n \rightarrow \infty} \Pr(\lambda(n) = \sum_{i=1}^n \lambda_i(n) = cf(n) \text{ is feasible}) = 1; \text{ and}$$

$$\liminf_{n \rightarrow \infty} \Pr(\lambda(n) = \sum_{i=1}^n \lambda_i(n) = c'f(n) \text{ is feasible}) < 1. \quad (1)$$

3. MAIN RESULTS

This section summarizes the main contributions of this paper. We first describe the upper bound capacity in MRWN.

- The Upper Bound:

The upper bound of the capacity in MRWN is given by the following theorem.

Theorem 3.1 *In a dense MRWN with uniform distribution of nodes and $s(n)$ source-destination pairs, the per source upper bound of the throughput capacity is given by*

$$C_{upper} = O \left(\min \left\{ \frac{\sqrt{\frac{n}{\log n}}}{s(n)}, 1 \right\} \right)^1. \quad (2)$$

If the number of source-destination pairs is larger than a threshold, i.e., $s(n) = \Omega \left(\sqrt{\frac{n}{\log n}} \right)$, then the upper bound capacity is dominated by this value. When $s(n)$ increases in the network while the number of relays decreases as $n - s(n)$, then the upper bound capacity decreases monotonically with $s(n)$. The main reason is the fact that there

¹Given two functions f and g , we say that: 1) $f(n) = O(g(n))$ if there exists a constant c and integer N such that $f(n) \leq cg(n)$ for $n > N$. 2) $f(n) = o(g(n))$ if $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$. 3) $f(n) = \Omega(g(n))$ if $g(n) = O(f(n))$. 4) $f(n) = \omega(g(n))$ if $g(n) = o(f(n))$. 5) $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $g(n) = O(f(n))$.

are not enough relays to facilitate the transportation of this information between all source-destination pairs and the lack of relays creates a bottleneck in the network. On the other hand, when $s(n)$ is smaller than a threshold, i.e., $s(n) = O \left(\sqrt{\frac{n}{\log n}} \right)$, then there are many relays

in the network to transport information for all source-destination pairs. Under this condition, it is clear that all source-destination pairs can transmit their information in parallel and the upper bound is 1. This result also indicates that when there is a large number of relays in the network, simple point-to-point communication does not utilize the full capability of the network. Under such conditions, it may be useful to develop cooperative techniques between nodes such that the network is able to fully take advantage of the relays in the network. Notice that when $s(n) = \Theta(n)$, the upper bound becomes $\Theta \left(\frac{1}{\sqrt{n \log n}} \right)$ which is similar to the result given in [1]. Under this condition, there will be no relay in the network and there are n simultaneous unicast sessions in the network.

- The lower Bound:

The achievable lower bound of the capacity is given by the following theorem.

Theorem 3.2 *In a dense MRWN with uniform node distribution, we assume there are $s(n)$ source-destination pairs. There exists a routing scheme which can lead to the following achievable throughput capacity for MRWN as*

$$C_{Lower} = \begin{cases} \Omega(1), & s(n) = \Theta(1) \\ \Omega \left(\frac{\log \log s(n)}{\log s(n)} \right), & \Omega(1) = s(n) = O(S_1) \\ \Omega \left(\frac{\sqrt{\frac{n}{\log n}}}{s(n)} \right), & s(n) = \Omega(S_1) \end{cases} \quad (3)$$

where $S_1 = \Theta \left(\sqrt{\frac{n}{\log n} \frac{\log n}{\log \log n}} \right)^2$.

The result indicates that for two regions of $s(n) = \Theta(1)$ and $s(n) = \Omega(S_1)$, the achievable lower bound is the same as the upper bound. However, when $\Omega(1) = s(n) = O(S_1)$, there exists a gap between the lower and upper bounds of the capacity. The reason behind this gap is in the randomness of the node distribution and the random selection of source-destination pairs.

The results of Theorems 3.1 and 3.2 are shown in Figure 2.

4. UPPER BOUND

In this section, we prove the upper bound of the capacity for the MRWN. For the wireless networks, we use the concept of *sparsity cut*, which is defined by Liu et al. [10], instead of min-cut, to take into account the differences between wired and wireless links.

²Note that this particular representation of S_1 is intentional in order to easier find the common regions between relay traffic and source-destination traffic in each cell later on.

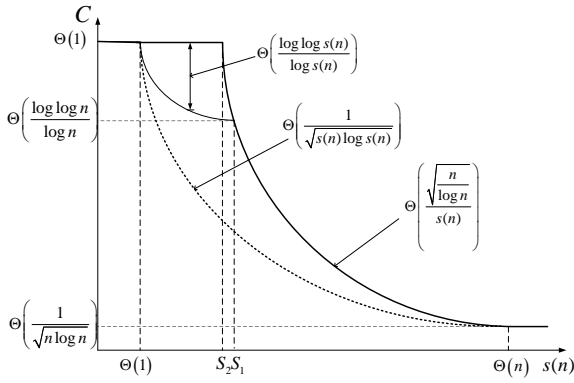


Figure 2: The upper and lower bound of the capacity in *MRWN* along with capacity of networks with no relays. The bold line is the upper bound of the capacity and the thin line is the lower bound of the capacity. In the regions of $(1, \Theta(1))$ and $[\Theta(S_1), n]$, the upper and lower bounds are tight. In this figure, $S_1 = \Theta\left(\sqrt{\frac{n \log n}{\log \log n}}\right)$, $S_2 = \Theta\left(\sqrt{\frac{n}{\log n}}\right)$.

Definition 4.1 (Sparsity Cut:) A sparsity cut for a random network is defined as a cut induced by the line segment with the minimum length that separates the region into two equal area subregions (see Fig 3). The cut capacity is defined as the transmission bandwidth W multiplied by the maximum possible number of simultaneous transmissions across the cut. This cut capacity is the information rate that the nodes from one side of the cut can deliver to the nodes at the other side. The cut length l_Γ is defined as the length of the cut line segment in 2-D space. In another word, sparsity cut can be seen for random geometric graph (RGG) similar to min-cut concept in graph theory.

From [1], we know that the disks centered at each receiver are disjoint and have radius of $\frac{\Delta r(n)}{2}$. By assuming the length of the sparsity cut as l_Γ , the following lemma provides the sparsity cut capacity which was originally proved in [10].

Lemma 4.2 The capacity of the cut Γ for a 2D region has an upper bound of $\frac{K_1 l_\Gamma W}{r(n)}$, where $K_1 = \max\left\{\frac{16}{\pi \Delta^2}, \frac{\sqrt{3}}{\Delta}\right\}$, and W is the link rate.

Since the network area is assumed to be 1, we have $l_\Gamma = \Theta(1)$. To guarantee the connectivity [1], $r(n)$ is chosen as $K_2 \sqrt{\frac{\log n}{n}}$. Thus, the upper bound of the cut capacity is $C_\Gamma = \frac{K_1 W}{K_2 \sqrt{\frac{\log n}{n}}} = \Theta\left(\sqrt{\frac{n}{\log n}}\right)$.

We use Lemma 4.2 to prove *Theorem 3.1* as described below.

Proof of Theorem 3.1: Lemma 4.2 shows the upper bound of the aggregate capacity of the network is $\Theta\left(\sqrt{\frac{n}{\log n}}\right)$. Since there are only $s(n)$ source-destination pairs in the network, it is clear that this throughput can be divided between these $s(n)$ source-destination pairs equally.

Therefore as long as $s(n)$ is larger than $\Theta\left(\sqrt{\frac{n}{\log n}}\right)$, per

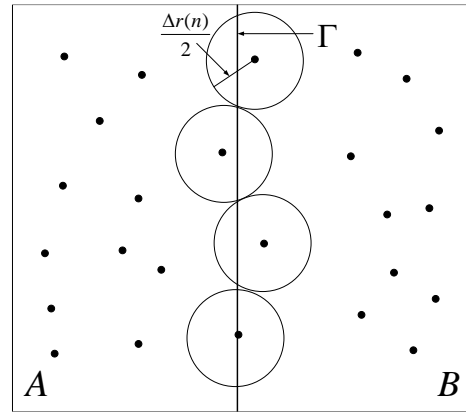


Figure 3: Γ is the sparsity cut which separates the network into A and B equal areas. The figure demonstrates four disjoint disks with radius $\frac{\Delta r(n)}{2}$ across the sparsity cut.

source throughput capacity is upper bounded as $\Theta\left(\frac{\sqrt{\frac{n}{\log n}}}{s(n)}\right)$.

However, when the number of source-destination pairs is less than $\Theta\left(\sqrt{\frac{n}{\log n}}\right)$, then there are plenty of capacity in the network such that each source can continuously transmit its data to destination. Under this condition, the per source throughput capacity is upper bounded as $\Theta(1)$.

Note that we consider point-to-point communications for this analysis. This result implies that when there are many relays in the network with few source-destination pairs, more advance communication schemes such as cooperative communications are appropriate to take advantage of relays in the network.

5. LOWER BOUND

In this section, the proof for the achievable lower bound of the throughput capacity is presented. We first describe the access scheme in the *MRWN*, then we present the "straight line" routing scheme. At the end, the achievable lower bound of *Theorem 3.2* is proven.

5.1 Access Scheme

We first divide the network into cells whose length is $s_n = K_3 \sqrt{\frac{\log n}{n}}$ to guarantee the connectivity between cells in the network. The cells are divided into M groups and in each time slot, only one group of cells are activated as shown in Fig. 4. The cell separation is such that successful communication based on protocol model is guaranteed when the cells in one group are activated simultaneously. The value of M is derived in the following lemma which was originally proposed in [11].

Lemma 5.1 There exists a positive integer $M = K_4(1 + \Delta)^2$ such that if we divide the network into M non interfering groups, then all the cells in a group can communicate every M time slots.

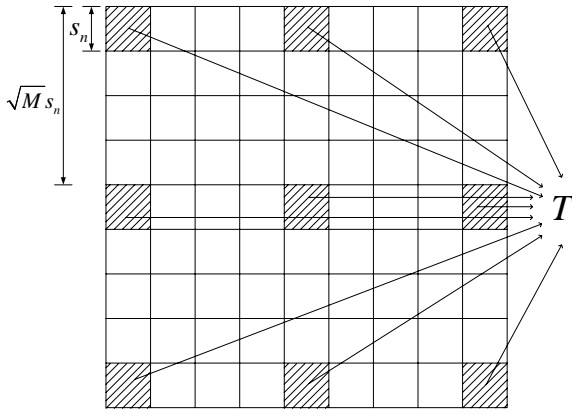


Figure 4: The shaded area cells have concurrent transmissions. $s_n = K_3 \sqrt{\frac{\log n}{n}}$ is the length of each cell and M is the number of the non interference groups.

5.2 Routing Scheme

The routing scheme is "straight line" routing that was described originally in [11]. In this routing scheme, for each source, we randomly and uniformly pick a location in the network and choose the closest node to this location as the destination for the source. The routing trajectory is a straight line L_i from the source node to this destination. Then the packets traverse from each source to the destination in a multihop fashion passing through all the cells that cross L_i .

5.3 Traffic In Each Cell

5.3.1 preliminaries

The achievable lower bound capacity is directly related to the number of lines passing through each cell. This achievable rate is proportional to the inverse of the number of lines passing through each cell.

There are two types of traffics in a cell. One type of traffic is caused by relays in the cell and the other one is caused by the sources and destinations in the cell. Unlike the cases in [1, 12] where the traffic is dominated by the latter traffic, in *MRWN* this assumption is not correct. In order to compute these two traffic types in *MRWN*, we first present the bin-ball problem lemma [8] and the Markov's and Chebyshev's inequalities [13] without any proof.

Lemma 5.2 Bin-Ball problem:

Let $N(m, n)$ be the random variable that counts the maximum number of balls in any bin when we throw m balls independently and uniformly at random into n bins. Then

$$N(m, n) = \begin{cases} \Theta\left(\frac{\log n}{\log \frac{n}{m}}\right), & \text{if } m < \frac{n}{\log n}, \\ \Theta\left(\frac{\log n}{\log \frac{n \log n}{m}}\right), & \text{if } \frac{n}{\text{polylog } n} \leq m \ll n \log n, \\ \Theta(\log n), & \text{if } m = c \cdot n \log n \\ & \text{for some constant } c, \\ \Theta\left(\frac{m}{n}\right), & \text{if } m \gg n \log n. \end{cases} \quad (4)$$

Lemma 5.3 Markov Inequality:

If X is any random variable and $a > 0$, then

$$\Pr(|X| \geq a) \leq \frac{E(|X|)}{a} \quad (5)$$

Lemma 5.4 Chebyshev Inequality:

Let X be a random variable with mean and standard deviation of μ_x and σ_x respectively. Then

$$\Pr(|X - \mu_x| \geq \alpha_x) \leq \frac{\sigma_x^2}{\alpha_x^2} \quad (6)$$

for any any $\alpha_x > 0$.

5.3.2 The Traffic Caused by Sources and Destinations

The traffic generated by the sources and the destinations in a cell is given by the following theorem.

Theorem 5.5 In *MRWN*, the maximum number of traffic T_{SD} caused by the sources and destinations in a cell is

$$T_{SD} = \begin{cases} \Theta\left(\frac{\log\left(\frac{n}{\log n}\right) \log s(n)}{\log\left(\frac{n}{s(n)}\right) \log \log s(n)}\right), & s(n) = O\left(\frac{n}{\log^2 n}\right) \\ \Theta\left(\frac{\log\left(\frac{n}{\log n}\right) \log s(n)}{\log\left(\frac{\frac{n}{\log n} \log\left(\frac{n}{\log n}\right)}{s(n)}\right) \log \log s(n)}\right), & \Omega\left(\frac{n}{\log^2 n}\right) \\ & = s(n) = o(n) \\ \Theta\left(\log\left(\frac{n}{\log n}\right) \frac{\log s(n)}{\log \log s(n)}\right), & s(n) = \Theta(n) \end{cases} \quad (7)$$

Proof: Let's denote the number of traffic caused by sources in each cell as T_S and the number of traffic caused by destinations in each cell as T_D . Clearly,

$$T_{SD} \leq \max\{T_S + T_D\} \leq \max\{T_S\} + \max\{T_D\} \quad (8)$$

Therefore, we need to compute both $\max\{T_S\}$ and $\max\{T_D\}$.

- Maximum Traffic Caused by Sources in Each Cell:

In each cell, since one source can only contribute to one flow, the maximum traffic caused by the sources in one cell is equal to the maximum number of sources in that cell. Since the side length of each cell is $K_3 \sqrt{\frac{\log n}{n}}$, the total number of cells in the network is $\frac{n}{K_3^2 \log n}$. We can apply the classical bin-ball problem by considering sources as balls and the cells as bins. By using Lemma 5.2, w.h.p.,

the maximum number of sources in each cell is

$$\max \{T_S\} = \begin{cases} \Theta \left(\frac{\log \left(\frac{n}{\log n} \right)}{\log \left(\frac{n}{s(n)} \right)} \right), & s(n) = O \left(\frac{n}{\log^2 n} \right) \\ \Theta \left(\frac{\log \left(\frac{n}{\log n} \right)}{\log \left(\frac{\frac{n}{\log n} \log \left(\frac{n}{\log n} \right)}{s(n)} \right)} \right), & \Omega \left(\frac{n}{\log^2 n} \right) = s(n) = o(n) \\ \Theta \left(\log \left(\frac{n}{\log n} \right) \right), & s(n) = \Theta(n) \end{cases} \quad (9)$$

In this derivation, the range of $s(n)$ is computed by assuming $\log \frac{n}{\log n} \cong \log n$.

- **Maximum Traffic Caused by Destinations in Each Cell:** Due to the randomness of the source-destination selection, there may exist several sources which have the same destination. Now, let's fit this problem into the bin-ball problem again. The balls and bins represent the sources and the destinations respectively. From *Lemma 5.2*, it is clear that there are at most $\Theta \left(\frac{\log s(n)}{\log \log s(n)} \right)$ sources w.h.p. for each destination³. Note that the maximum number of destinations in each cell is the same as that of sources or equivalently $\max \{T_S\}$ given by (9). Thus $\max \{T_D\}$ is given by

$$\max \{T_D\} = \max \{T_S\} \cdot \Theta \left(\frac{\log s(n)}{\log \log s(n)} \right). \quad (10)$$

Therefore, we arrive at

$$\begin{aligned} T_{SD} &\leq \max \{T_S\} + \max \{T_D\}, \\ &= \max \{T_S\} + \max \{T_S\} \Theta \left(\frac{\log s(n)}{\log \log s(n)} \right), \\ &= (\max \{T_S\} + 1) \Theta \left(\frac{\log s(n)}{\log \log s(n)} \right). \end{aligned} \quad (11)$$

By combining (9) and (11), the theorem follows.

5.3.3 Traffic Caused by Relays

In this section, we compute the traffic caused by relays. First, we introduce a lemma from [11].

Lemma 5.6 *In a network with uniform distribution of nodes, there exists a positive constant K_5 such that for every line L_i and cell C_j ,*

$$\Pr(\text{Line } L_i \text{ intersects } C_j) \leq K_5 \sqrt{\frac{\log n}{n}} \quad (12)$$

The maximum traffic caused by relays is given by the following theorem.

³Note that we used the second line in the equation of *Lemma 5.2* because the number of sources and destination is equal to $s(n)$.

Theorem 5.7 *In MRWN, the maximum relay traffic T_R in a cell is given by*

$$\max \{T_R\} = \begin{cases} \Theta(1), & \Omega(1) = s(n) = o \left(\sqrt{\frac{n}{\log n}} \right) \\ \Theta(f(n)), & s(n) = \Theta \left(\sqrt{\frac{n}{\log n}} \right) \\ \Theta \left(s(n) \sqrt{\frac{\log n}{n}} \right), & s(n) = \omega \left(\sqrt{\frac{n}{\log n}} \right) \end{cases} \quad (13)$$

where $f(n)$ can be any function of n that fulfills the condition $\lim_{n \rightarrow \infty} f(n) = \infty$.

proof: In this proof, we divide $s(n)$ into three regions which are $s(n) = \Theta(1)$, $\omega(1) = s(n) = O \left(\sqrt{\frac{n}{\log n}} \right)$ and $s(n) = \omega \left(\sqrt{\frac{n}{\log n}} \right)$.

- **Case of $s(n) = \Theta(1)$:** In this region, $s(n)$ is a positive constant. Clearly, the traffic caused by relays is at most equal to $s(n)$. Thus, $\max \{T_R\} = \Theta(1)$.

- **Case of $\omega(1) = s(n) = O \left(\sqrt{\frac{n}{\log n}} \right)$:**

In this region, we will prove that when $\omega(1) = s(n) = o \left(\sqrt{\frac{n}{\log n}} \right)$, then $\max \{T_R\} = \Theta(1)$. When $s(n) = \Theta \left(\sqrt{\frac{n}{\log n}} \right)$, then $\max \{T_R\} = \Theta(f(n))$, where $f(n)$ is defined above.

We first introduce the following lemma.

Lemma 5.8 *In MRWN, when the "straight line" routing scheme is used and k is a positive constant, we have*

$$\begin{aligned} &\Pr(k \text{ lines intersecting any cell } C_j) \\ &\leq \binom{s(n)}{k} \left(K_5 \sqrt{\frac{\log n}{n}} \right)^k \left(1 - K_5 \sqrt{\frac{\log n}{n}} \right)^{s(n)-k} \end{aligned} \quad (14)$$

This lemma is proved in the Appendix 7.

For $s(n) = \omega(1)$ in conjunction with *Lemma 5.8* and large n , we arrive at

$$\begin{aligned} &\Pr(k \text{ lines intersecting any cell } C_j) \\ &\leq \lim_{s(n) \rightarrow \infty} \binom{s(n)}{k} \left(K_5 \sqrt{\frac{\log n}{n}} \right)^k \left(1 - K_5 \sqrt{\frac{\log n}{n}} \right)^{s(n)-k}, \\ &= \lim_{s(n) \rightarrow \infty} \binom{s(n)}{k} \left(\frac{K_5 s(n) \sqrt{\frac{\log n}{n}}}{s(n)} \right)^k \left(1 - \frac{K_5 s(n) \sqrt{\frac{\log n}{n}}}{s(n)} \right)^{s(n)-k}, \\ &\leq \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}} \right)^k}{k!} \exp \left(-K_5 s(n) \sqrt{\frac{\log n}{n}} \right). \end{aligned} \quad (15)$$

The last line of equation is derived by considering the fact that it is Poisson distribution. Let's define K_6 as the maximum number of lines passing through each cell, then one can obtain

$$\begin{aligned}
& \Pr(k \leq K_6) \\
&= \sum_{k=0}^{K_6} \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)^k}{k!} \exp\left(-K_5 s(n) \sqrt{\frac{\log n}{n}}\right) \\
&\stackrel{(a)}{=} \frac{\sum_{k=0}^{K_6} \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)^k}{k!}}{\sum_{i=0}^{\infty} \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)^i}{i!}} \\
&= \frac{\sum_{k=0}^{\infty} \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)^k}{k!} - \sum_{k=K_6+1}^{\infty} \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)^k}{k!}}{\sum_{i=0}^{\infty} \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)^i}{i!}} \\
&= 1 - \frac{\sum_{k=K_6+1}^{\infty} \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)^k}{k!}}{\sum_{i=0}^{\infty} \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)^i}{i!}} \\
&\stackrel{(b)}{\geq} 1 - \frac{\sum_{k=K_6+1}^{\infty} \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)^k}{K_6+1}}{\exp\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)} \\
&= 1 - \frac{\frac{1}{K_6+1} \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)^{K_6+1}}{1 - K_5 s(n) \sqrt{\frac{\log n}{n}}}}{\exp\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)} \tag{16}
\end{aligned}$$

(a) and (b) are due to the Maclaurin series that $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. Now there are two cases with respect to $s(n)$.

– When $s(n) = o\left(\sqrt{\frac{n}{\log n}}\right)$:

Since K_6 is a constant value and as n tends to infinity, from (16) it can be concluded that

$$\begin{aligned}
\Pr(k \leq K_6) &= \lim_{n \rightarrow \infty} 1 - \frac{\frac{1}{K_6+1} \frac{\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)^{K_6+1}}{1 - K_5 s(n) \sqrt{\frac{\log n}{n}}}}{\exp\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)} \\
&= 1, \tag{17}
\end{aligned}$$

which implies that $\max\{T_R\} \leq K_6$.

– When $s(n) = \Theta\left(\sqrt{\frac{n}{\log n}}\right)$:

In this case, (16) cannot be used any more. However, Markov inequality in *Lemma (5.3)* implies that

$$\Pr(T_R \geq K_6) \leq \frac{E(T_R)}{K_6},$$

$$\begin{aligned}
&= \frac{E\left(K_5 s(n) \sqrt{\frac{\log n}{n}}\right)}{K_6}, \\
&= \frac{K_7}{K_6}, \tag{18}
\end{aligned}$$

where K_7 is a positive constant. From (18), we can see that if $K_6 = f(n)$ where $f(n)$ is an arbitrary function of n satisfying $\lim_{n \rightarrow \infty} f(n) = \infty$, then as n goes to infinity, we can have

$$\Pr(T_R \geq K_6) = 0. \tag{19}$$

Thus $\max\{T_R\} \leq f(n)$.

• Case of $s(n) = \omega\left(\sqrt{\frac{n}{\log n}}\right)$:

In this region of $s(n)$, we will show that $\max\{T_R\} = O\left(s(n) \sqrt{\frac{\log n}{n}}\right)$.

Let's assume $\lim_{n \rightarrow \infty} \frac{\alpha}{s(n) K_5 \sqrt{\frac{\log n}{n}}} = \varepsilon$ where ε is a positive constant. By utilizing Chebychev inequality in *Lemma 5.4*, it can be shown that

$$\begin{aligned}
&\Pr\left(\left|T_R - s(n) K_5 \sqrt{\frac{\log n}{n}}\right| \geq \sqrt{\alpha s(n) \sqrt{\frac{\log n}{n}}}\right) \\
&\leq \frac{s(n) K_5 \sqrt{\frac{\log n}{n}} \left(1 - K_5 \sqrt{\frac{\log n}{n}}\right)}{\alpha s(n) K_5 \sqrt{\frac{\log n}{n}}} \\
&= \frac{1 - K_5 \sqrt{\frac{\log n}{n}}}{\alpha} \tag{20}
\end{aligned}$$

which goes to zero as n tends to infinity and we used the fact that the random variable has Poisson distribution. Thus $\max\{T_R\} \leq s(n) K_5 \sqrt{\frac{\log n}{n}}$. Notice that from this result we can find that when $s(n) = n$, $T_R = O(\sqrt{n \log n})$ which is the same as the result in [1]. Note that under this condition, the traffic is dominated by relays and source or destination traffic in each cell are simply negligible.

5.4 Achievable Throughput

The total traffic in any cell can be obtained by using the results shown in Sections 5.3.2 and 5.3.3 which is summarized here.

$$\begin{aligned}
T_{\text{total}} &\leq \max\{T_{\text{SD}}\} + \max\{T_R\} \\
&= \begin{cases} O(1), & s(n) = \Theta(1) \\ O\left(\frac{\log\left(\frac{n}{\log n}\right)}{\log\left(\frac{n}{s(n)}\right)} \frac{\log s(n)}{\log \log s(n)}\right), & \Omega(1) = s(n) = O(S_1) \\ O\left(s(n) \sqrt{\frac{\log n}{n}}\right), & s(n) = \Omega(S_1) \end{cases} \tag{21}
\end{aligned}$$

where $S_1 = \Theta\left(\sqrt{\frac{n}{\log n} \frac{\log n}{\log \log n}}\right)$. From *Lemma 5.1*, we can find that there exists a transmission scheme such that in every $M = K_4(1+\Delta)^2$ slots, each cell can get one slot to send packets at a rate W bits/second. Thus, the rate for each cell is $\frac{W}{K_4(1+\Delta)^2}$. From Eq.(21), each cell can send packets at a rate equal to T_{total} with probability one as n goes to infinity. Therefore, the maximum achievable throughput C_{Lower} for each source should satisfy

$$C_{\text{Lower}} T_{\text{total}} = \frac{W}{K_4(1+\Delta)^2}. \quad (22)$$

Hence

$$C_{\text{Lower}} = \frac{\frac{W}{K_4(1+\Delta)^2}}{T_{\text{total}}} \geq \frac{\frac{W}{K_4(1+\Delta)^2}}{\max\{T_{\text{SD}}\} + \max\{T_{\text{R}}\}} \quad (23)$$

Theorem 3.2 follows immediately. Note that in derivation of second line in *Theorem 3.2*, we ignore the term $\frac{\log\left(\frac{n}{\log n}\right)}{\log\left(\frac{\log n}{s(n)}\right)}$ because this term is asymptotically equal to $\Theta(1)$.

6. DISCUSSION

6.1 Gains from Multihop Relays

From the definition of the network model, we know that the relays in *MRWN* only utilize the decode-and-forward operation to help the source to transmit information without any cooperation. Even with no cooperation, the network with relays provide order throughput gain compared to the case of no relays in the network. The capacity of this type of network with no relay, $s(n)$ source-destination pairs and by only using plain multihop routing scheme is given in [1], which is $\frac{1}{\sqrt{s(n)\log s(n)}}$ ⁴. Therefore, the gains G from the relays in the network is given by

$$G = \frac{C_{\text{Lower}}}{\frac{1}{\sqrt{s(n)\log s(n)}}} \quad (24)$$

$$\begin{cases} \Omega\left(\sqrt{s(n)\log s(n)}\right), & s(n) = \Theta(1) \\ \Omega\left(\log \log s(n) \sqrt{\frac{s(n)}{\log s(n)}}\right), & \omega(1) = s(n) = O(S_1) \\ \Omega\left(\sqrt{\frac{n}{s(n)} \frac{\log s(n)}{\log n}}\right), & s(n) = \Omega(S_1) \end{cases}$$

The first interesting observation from (24) is the fact that by increasing $s(n)$, the gain G in some regions increases while in other regions decreases. More specifically, when $s(n) = \Omega(S_1)$ then the gain G decreases with the increase of $s(n)$ and when $s(n) = O(S_1)$, G increases with the increase in $s(n)$.

When $s(n) = \Theta(S_1)$, the gain G obtains the maximum value which means in order to get the optimal gain from relays, we need to make the traffic caused by the relays be comparable to the traffic caused by the sources and destinations

⁴Note that unlike [1] that considers n source-destination pairs, we assume $s(n)$ source-destination pairs in order to compare it with our technique

in a cell. The reason of this interesting result is that in the case that $s(n) = \Omega(S_1)$, as the increase of $s(n)$, the number of relays becomes smaller and smaller, so the gain caused by relays decreases. While in the case that $s(n) = O(S_1)$, $s(n)$ is so small that not all of the relays can work in parallel, so as the increase of $s(n)$, more and more relays help in the network for transportation of information, thus the gain obtained from relays increases with the increase of $s(n)$. When all the relays helps sources in delivery of data and $s(n)$ is not too large, the gain G achieves its optimal value.

We know that in *MRWN*, the transmission scheme is plain multihop routing without any cooperative scheme. One important question is "why do we achieve such order gains in throughput capacity by using simple routing and relays?" This gain mainly comes from the increase in the number of concurrent transmissions by increasing the number of relays. When the total network area is constant and we increase the number of relays, the distance between nodes decreases which can result in decrease in transmission range. By decreasing the transmission range, one can increase the number of concurrent transmissions in the network which is much larger than the increase in the number of hops due to reduction in transmission range. Next figure demonstrates the gain from relays in the networks compared to the case of no relay utilizing multihop point-to-point communications.

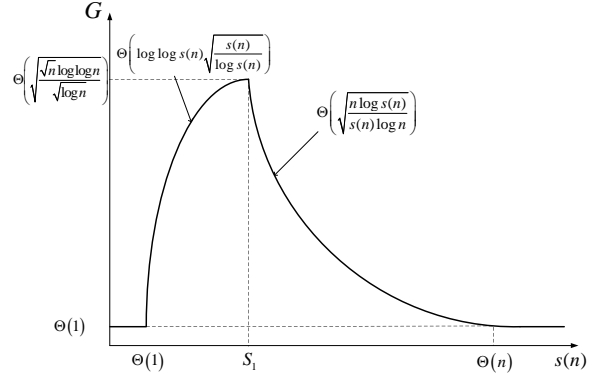


Figure 5: The gain in *MRWN* compared to that of networks with no relays. In both schemes, simple point-to-point communication protocol is utilized.

6.2 Gap Between Lower and Upper Bounds

Another important observation is that, from Fig. 2, there is a gap between the lower and upper bound of the capacity for *MRWN* in the region of $s(n) = [1, S_1]$. In the region of $s(n) = [S_1, n]$, it is clear that the main restriction of the network is from the traffic caused by relays, while in the region of $s(n) = [1, S_1]$, the main bottleneck of the network comes from the traffic caused by the sources and destinations in each cell. From the bin-ball problem, we observe that this gap comes from the combination of the maximum number of destinations in each cell and the maximum number of sources for each destination. The upper bound of the traffic in each cell is caused by these two values. These maximum numbers are much larger than the mean number of destinations in each cell or mean number of sources for each destination. This large difference is caused by the randomness of the net-

work. However, if we let each cell have the same number of sources and destinations, and the traffic becomes the permutation traffic which means that each destination only has one source, then this gap will disappear.

This observation gives us a hint that randomness can lead to the loss of the capacity in the real network. In cells with dominant traffic dictate the throughput capacity for that network. In order to achieve higher capacity, we should do more averaging and remove the randomness.

7. CONCLUSION

In this work, we introduce the *Multihop Relay Wireless Networks*. In this network model, there are only $s(n)$ out of n nodes generating information and other nodes act as relays. Additionally, the transmission scheme in this network is multihop and plain routing rather than any cooperative scheme. Under this network assumption, we compute upper and lower bounds of the throughput capacity. From the results, we find that for different values of source-destination pairs $s(n)$, the behaviour of the network is different. When $s(n)$ is small, the constraint of the network is $s(n)$ or equivalently, it is caused by the traffic from the sources and the destinations. While $s(n)$ is large, the constraint of the network comes from the maximum sustainable capacity that relays can transport in the network or equivalently, it is caused by the relay traffic in the network. Interestingly, even though we only use the relays in multihop fashion without any cooperative scheme, we can still achieve some order gain from the relays, which comes from the increase in the number of concurrent transmissions caused by increasing the number of relays. When the amount of traffic caused by the relays is comparable to the traffic load caused by the sources and destinations in a cell, the gain from the relays can be maximized. Furthermore, from the derivation of the capacity, there exists a gap between the upper and lower bound of the capacity, which can be understood as the loss due to the randomness of the network. By removing the randomness of the network, this gap can be closed.

Acknowledgement

This research was partially sponsored by the U.S. Army Research Laboratory under the Network Science Collaborative Technology Alliance, Agreement Number W911NF-09-0053, by the Army Research Office under agreement number W911NF-05-1-0246, by the National Science Foundation under grant CCF-0729230, and by the Baskin Chair of Computer Engineering. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

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, March 2000.
- [2] Z. Wang, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves, "A unifying perspective on the capacity of

wireless ad hoc networks," in *Infocom*, April 2008, pp. 753–761.

- [3] L.-L. Xie and P.R.Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation," *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 748–767, May 2004.
- [4] A. Ozgur, O. Lévêque, and D. N.C.Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 3549–3572, October 2007.
- [5] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2004.
- [6] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," in *INFOCOM*, June 2002, pp. 1577–1586.
- [7] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IT*, vol. 51, no. 9, pp. 3037–3063, September 2005.
- [8] B. Liu, D. Towsley, and A. Swami, "Data gathering capacity of large scale multihop wireless networks," in *Mobihoc*, 2008.
- [9] E. Perevalov and D. S. Rick S.Blum, "Capacity of clustered ad hoc networks: How large is large," *IEEE Transactions on Communication*, vol. 54, no. 9, pp. 1672–1681, September 2006.
- [10] J. Liu, D. Goeckel, and D. Towsley, "The throughput order of ad hoc networks employing network coding and broadcasting," in *MILCOM*, 2006.
- [11] F. Xue and P. R. Kumar, *Scaling Laws for Ad Hoc Wireless Networks: An Information Theoretic Approach*. NOW Publishers, 2006.
- [12] S. R.Kulkarni and P. Viswanath, "A deterministic approach to throughput scaling in wireless networks," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1041–1049, June 2004.
- [13] B. Motwani and P. Raghavan, *Randomized Algorithms*. Cambridge University Press, 1995.

Appendix

The Proof of Lemma 5.8

We assume that $\Pr(\text{Line } L_i \text{ intersects } C_j) = p$. Since the events that different line L_i intersects C_j are independent, then

$$\begin{aligned} & \Pr(\text{There are } k \text{ lines intersecting any cell } C_j) \\ &= \binom{s(n)}{k} (p)^k (1-p)^{s(n)-k} \end{aligned} \quad (25)$$

By doing the derivative of (25), one arrives at the maximum value of (25) given by $p = \frac{k}{s(n)} = p^*$. when $p < p^*$, then this probability increases with increase in p . Let's assume $K_5 \sqrt{\frac{\log n}{n}} = p^\dagger$. Since $s(n) = O\left(\sqrt{\frac{n}{\log n}}\right)$, it is clear that $p^* \geq kK_6 \sqrt{\frac{\log n}{n}}$, where K_6 is a positive constant. By allowing k large enough, we obtain $p^\dagger \leq p^*$. By using Lemma 5.6, it is clear that $p \leq p^\dagger$. Due to monotonic increase of $\Pr(\text{There are } k \text{ lines intersecting any cell } C_j)$ with the increase in p , the result follows.