

Data Gathering Capacity of Large Scale Multihop Wireless Networks

Benyuan Liu

Department of Computer Science
University of Massachusetts at Lowell
Lowell, MA 01854

Don Towsley

Department of Computer Science
University of Massachusetts at Amherst
Amherst, MA 01002

Ananthram Swami

Army Research Laboratory
Adelphi, MD 20783

Abstract

This paper studies the scaling laws of the data gathering capacity of large scale multihop wireless networks. Unlike the data communication paradigms studied in previous research, for example, the many-to-many, many-to-one, broadcast, and multicast paradigms, the data gathering capacity concerns the per source node throughput in a network where a subset of nodes send data to some designated destinations while other nodes serve as relays. This some-to-some communication paradigm is commonplace in many wireless networks, for example, wireless mesh networks and wireless sensor networks, and in some cases perhaps more prevalent than the other paradigms. We first derive the upper and constructive lower bounds for the data gathering capacity, and then examine their design and performance implications. Our results show that the data gathering capacity is constrained by different factors in several different scaling regimes of the number of source and destination nodes, exhibiting distinct scaling laws in those regimes. This work fills a gap in our understanding of the capacity of various communication paradigms, and can lead to better network planning and performance for data gathering wireless network applications.

1 Introduction

Ad hoc networking has become a critical technology for wireless nodes to communicate with each other in environments where there is no network infrastructure. The ad hoc network architecture will be widely used in a variety of applications including wireless sensor networks, disaster recovery, and military on-the-field communications.

Recently there has been increasing interest in studying the scaling laws of the capacity of large scale wireless ad

hoc networks. The capacity of a wireless network is a fundamental performance metric that characterizes the long-term throughput of each source node to its destination. The scaling laws of the capacity is dictated by many different aspects of the network, for example, network architecture, radio interference model, node communication paradigm, traffic locality, node mobility, power and bandwidth constraints. A solid understanding of the capacity scaling laws is important for the design, planning, and performance characterization for various network scenarios.

Many of the previous studies of the wireless network capacity have focused on the many-to-many node communication paradigm first considered in [7], where every node is a data source, and the destination of a source node is randomly chosen within the network. Recently, the network capacity of several other communication paradigms have been investigated. These include the one-to-one relay capacity, where there is only one source-destination pair while other nodes serve as relays [4]; the many-to-one capacity, where all of the nodes send data to one sink node [13, 5, 3]; the broadcast capacity, where one node broadcasts data to all of the other nodes in the network [19, 9, 14]; and the multicast capacity, where some nodes send data to other nodes in their respective multicast groups [8, 11, 18].

In this paper we consider a different communication paradigm that has not been studied before, namely, the data gathering paradigm. In this paradigm, not every node has data to send, and not every node can be a destination. A subset of the nodes send data to some designated destinations while other nodes without their own data serve as relays. This some-to-some communication paradigm is widely present in many application scenarios, and in some cases perhaps more prevalent than other paradigms. For example, in a wireless mesh network, only nodes that have data will send to selected access points or gateway nodes. Also, in many wireless sensor network applications, only

those sensors that have sensed interesting events have data to send, and the data sinks are usually chosen from some designated nodes in the network.

The goal of this work is to study the scaling laws of the data gathering capacity of a wireless network. Interesting questions arise. How does this capacity scale with the number of source nodes, the number of destination nodes, as well as the total number of nodes in the system? How does the capacity compare with those of other communication paradigms? What are the design and performance implications? Answers to these questions will fill an important gap on our understanding of the capacity for various communication paradigms, and lead to better network design and planning tools for a range of wireless network applications.

We obtain the upper and constructive lower bounds of the data gathering capacity. Based on the results we explore their design and performance implications. First, we find that the upper bound of the data gathering capacity is constrained by different factors and exhibits distinct scaling behaviors in different scaling regimes of the number of destination nodes. Specifically, consider a random wireless network of n nodes, among which there are $s(n)$ randomly selected source nodes and $d(n)$ randomly selected destination nodes. Each source node is randomly mapped to a destination node. If the number of randomly selected destination nodes is asymptotically less than $\sqrt{n/\log n}$, the capacity is constrained by the number of destination nodes. The network has enough transport capacity to saturate all of the destination nodes. Otherwise, the capacity is constrained by the transport capacity of the network. The scaling behaviors of the capacity is illustrated in Figure 1.

The characterizations of the capacity allows us to identify the bottleneck of the data gathering capacity for given network scenarios and choose proper approaches to improve the capacity. For example, when the capacity is constrained by the number of available destination nodes, it is beneficial to add extra destination nodes. On the other hand, if the capacity is constrained by the network transport capacity to relay the packets, adding more destination nodes will not result in substantial capacity improvement.

We devise a node transmission scheduling algorithm and a routing algorithm to obtain an constructive lower bound of the data gathering capacity. We find that in certain scaling regimes of the number of source and destination nodes, there is a gap between the constructive lower bound and the upper bound of the data gathering capacity. This gap is attributed to the random selection of the source and destination nodes among the nodes and the random mapping between them. Based on our analysis, in Section 5 we propose several mechanisms to improve the constructive lower bound and hence narrow the gap.

The remainder of the paper is organized as follows. In Section 2 we describe the network model. The main contri-

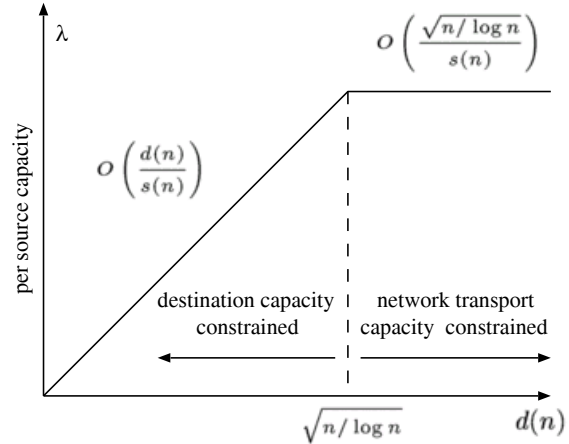


Figure 1. Upper bound of data gathering capacity in different scaling regimes of number of destination nodes, $d(n)$.

butions of this paper is summarized in Section 3. We derive the upper bound and constructive lower bound of the data gathering capacity in Section 4 and 5, respectively. We finally conclude the paper in Section 6.

2 Network Model

We consider a dense network where n wireless nodes are distributed in a unit area square uniformly at random. Not every node has data to send, and not every node can be a receiver. A subset of nodes are sources, which have data to send to some designated destination nodes. We assume that, among the n nodes, there are $s(n)$ randomly selected source nodes and $d(n)$ randomly selected destination nodes. Each source node is randomly mapped to a destination node. Without loss of generality, we consider the case where the number of source nodes is no less than the number of destination nodes, i.e., $s(n) \geq d(n)$.

Note that the above traffic model can also be used to model scenarios where a smaller number of source nodes send independent data to a larger number of destination nodes, i.e., $s(n) \leq d(n)$. This differs from multicast as the data to the destination nodes are independent, while in multicast the same data are sent to the destination nodes.

All of the nodes share a common wireless channel. Each node has a common transmission range r . For simplicity, we assume each node can transmit data at a rate of 1 bit/sec over the channel when scheduled. There is only a constant factor in the final results if the actual node transmission rate is W bits/sec. A time-division multiplexing (TDMA) scheme is used for scheduling data transmission over the wireless channel. Time is divided into fixed dura-

tion slots. In each time slot, certain nodes are scheduled to send data. A node cannot transmit and receive data simultaneously, and a node can only receive data from one other node at a time.

We adopt the Protocol Interference Model introduced in [7]. A transmission from node X_i is successfully received by node X_j if the following two conditions are satisfied: (i) Node X_j is within the transmission range of node X_i , i.e., $|X_i - X_j| \leq r$, where $|X_i - X_j|$ represents the distance between node X_i and node X_j in the plane; (ii) For every other node X_k that is simultaneously transmitting over the same channel, $|X_k - X_j| \geq (1 + \Delta)|X_i - X_j|$. This condition guarantees that for each successful transmission, there is a guard zone around the receiving node such that no nodes in the guard zone can transmit the same time. The radius of the guard zone is $(1 + \Delta)$ times the distance between the sender and receiver. The parameter Δ is a fixed positive constant that determines the size of the guard zone.

The protocol interference model tries to capture the behavior of existing protocols such as IEEE 802.11, and hence has important practical implications. A more realistic interference model is the physical model, where a transmission is successful if the signal-to-interference-noise ratio (SINR) is above a threshold. In this paper we focus on the protocol model but we expect similar results hold for the physical model.

We study the asymptotic capacity scaling law of the data gathering networks. An event is said to occur with high probability (w.h.p.) if its probability tends to 1 as $n \rightarrow \infty$. The per source node data gathering capacity of a wireless network is defined as follows.

Definition 1 Consider a wireless network where n nodes are distributed in a unit area square uniformly at random. Among the n nodes there are $s(n)$ randomly selected source nodes and $d(n)$ randomly selected destination nodes. Each source node is randomly mapped to a destination node. The throughput capacity per source node $\lambda(n, s(n), d(n))$ is the number of bits per second that every source node can transmit w.h.p. to its destination.

Throughout the paper we use $O(g(x))$, $\Omega(g(x))$ and $\Theta(g(x))$ to represent the asymptotic upper, lower and tight bounds of function $g(x)$.

In practice, the mapping between source and destination nodes may not be uniformly random. For example, if the goal is to send data to any of the data sinks (anycast), a sensor is more likely to send its data to the nearest data sink rather than to far away ones. In this case, the locality of the data traffic plays an important role in determining the data gathering capacity. Nevertheless, there are many application scenarios in which the uniformly random mapping model makes sense. A good example is the geographic hash table (GHT) mechanism proposed for data-centric storage

in sensor networks [17]. In this approach, sensor data is routed and stored at nodes rather uniformly across the network using a distributed hashing table system. Also, in other sensor network applications, independent users with different interests may query sensors at different locations, which can be approximated by the some-to-some communication model.

Useful Known Results of Maximum Occupancy.

In the construction of lower bound in Section 5, we will need to use the following known results on the maximum occupancy of randomly locations.

Suppose a total number of m nodes are distributed into n equal sized areas independently and uniformly at random. Let $M(m, n)$ denote the maximum number of nodes located in each area. We want to know the scaling behaviors of $M(m, n)$. This can be mapped to the well known ‘‘balls into bins’’ problem. The following maximum occupancy result is adapted from [16][Theorem 1] and [15][Lemma 2.13].

Lemma 1 Assume m nodes are placed into n equal sized areas independently and uniformly at random. Let $M(m, n)$ be the random variable that counts the maximum number of nodes in any area. With high probability (whp), we have

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

When $m \geq cn \log n$ for some constant c , the maximum number of nodes in each area is of the same order as the average number m/n . However, when $m \ll n \log n$, the maximum number in each area start to deviate from the average. In the special case where $m = n$, we have $M(m, n) = \log n / \log \log n$, which is much larger than the average one node per area.

3 Summary of Contributions

Theorem 1 Consider a wireless network of n nodes randomly located in a unit area square, and among the nodes there are $s(n)$ randomly selected source nodes and $d(n)$ randomly selected destination nodes. Each source node is randomly mapped to a destination node. Without loss of generality, we assume $s(n) \geq d(n)$. The per source node capacity is upper bounded by

$$\lambda(n, s(n), d(n)) = O\left(\min\left(\frac{d(n)}{s(n)}, \frac{\sqrt{n/\log n}}{s(n)}\right)\right). \quad (1)$$

Depending on the scaling regimes of the number of destination nodes, $d(n)$, the upper bound of data gathering capacity is constrained by different factors and exhibits different scaling laws, as illustrated in Figure 1.

- $d(n) = \omega(\sqrt{\frac{n}{\log n}})$.

When the number of destination nodes is asymptotically larger than $\sqrt{\frac{n}{\log n}}$, the upper bound becomes:

$$\lambda(n, s(n), d(n)) = O\left(\frac{\sqrt{n/\log n}}{s(n)}\right).$$

In this case the data gathering capacity is limited by the network transport capacity. In the special case where $s(n) = d(n) = \Theta(n)$, we obtain

$$\lambda(n, s(n), d(n)) = O\left(\frac{1}{\sqrt{n \log n}}\right),$$

which recovers the results in [7, 10].

- $d(n) = O(\sqrt{\frac{n}{\log n}})$.

When the number of destination nodes is asymptotically equal to or smaller than $\sqrt{\frac{n}{\log n}}$, the upper bound becomes:

$$\lambda(n, s(n), d(n)) = O\left(\frac{d(n)}{s(n)}\right).$$

In this case, the data gathering capacity is limited by the number of destination nodes. The network has enough transport capacity to saturate all of the destination nodes.

The characterization of the upper bound allows us to identify the bottleneck in the data gathering network for a given network scenario, based on which we can choose proper approaches to improve the capacity. Specifically, when the capacity is constrained by the number of available destination nodes, it is beneficial to add extra destination nodes. On the other hand, if the capacity is constrained by the network transport capacity to relay the packets, adding more destination nodes will not result in substantial capacity improvement. Instead, one can consider adding more relay nodes or other capacity improvement approaches, for example, exploiting node mobility [6], adding infrastructure [12], using coherent beam-forming [1] or multiple-input multiple-output (MIMO) techniques [2], etc.

We further obtain a constructive lower bound for the data gathering capacity of a random wireless network.

Theorem 2 *In the aforementioned random wireless network, there exists a scheduling and routing scheme such that the following per source capacity can be achieved.*

$$\lambda(n, s(n), d(n)) = \Omega\left(\min\left(\frac{1}{M(s(n), \sqrt{n/\log n})}, \frac{1}{M(s(n), d(n))M(d(n), \sqrt{n/\log n})}\right)\right) \quad (2)$$

where $M(m, n)$ is given in Lemma 1, denoting the maximum number of nodes located in each area when a total number of m nodes are distributed into n equal sized areas independently and uniformly at random.

For certain scaling regimes of $s(n)$ and $d(n)$, there is a gap between the constructive lower bound and the upper bound. This gap is attributed to the random selection of the source and destination nodes and the random mapping between them. Based on our analysis of the gap, in Section 5 we propose mechanisms to improve the constructive lower bound and hence narrow the gap.

4 Upper Bound on Data Gathering Capacity in Random Networks

The upper bound of the data gathering capacity is determined by three factors. These are: 1) sender capacity constraint, 2) receiver capacity constraint, and 3) network transport capacity constraint. In the following we will first derive the three capacity constraints, obtain the upper bound, and then examine its scaling behaviors in different regimes.

1. *Sender capacity constraint.* The sending rate of a source node over the wireless channel is 1 bit/sec. The per-source capacity cannot exceed this rate, i.e.,

$$\lambda(n, s(n), d(n)) \leq 1. \quad (3)$$

2. *Receiver capacity constraint.* There are $d(n)$ receivers, and at any time instant, a receiver can only receive data from one other node at 1 bit/sec. Since there are $s(n)$ source nodes, the per-source capacity is constrained by

$$\lambda(n, s(n), d(n)) \leq \frac{d(n)}{s(n)}. \quad (4)$$

3. *Network transport constraint.* The per-source throughput is constrained by the network transport capacity as packets are transported from the source to the destination in the network. The derivation of the network transport capacity per source node is similar to that in [7], with a change in the number of source nodes.

Under the protocol model, the receivers of successful transmissions cannot be within distance $\Delta r(n)$ of each

other. Each successful transmission thus occupies at least a disk of radius $\Delta r(n)/2$. The number of simultaneous transmissions in the unit area square is bounded from above by $\frac{4}{\pi \Delta^2 r(n)^2}$.

As the source and destination nodes are independently and uniformly distributed in the unit area disk, and each source node chooses the destination node uniformly at random, the mean path is $\Theta(1)$. Thus the mean number of hops taken by a packet is at least $c/r(n)$ for some $c > 0$. Since each source generates $\lambda(n, s(n), d(n))$ bits/sec, the aggregate data rate from all source nodes is at least $\frac{cs(n)\lambda(n, s(n), d(n))}{r(n)}$. We have

$$\frac{cs(n)\lambda(n, s(n), d(n))}{r(n)} \leq \frac{4}{\pi \Delta^2 r(n)^2}.$$

Thus,

$$\lambda(n, s(n), d(n)) \leq \frac{4}{c\pi \Delta^2 s(n)r(n)}.$$

It has been shown in [7] that $r(n) > \sqrt{\log n/n}$ is necessary to guarantee that the network is connected with high probability. Therefore,

$$\lambda(n, s(n), d(n)) \leq \frac{c' \sqrt{n/\log n}}{s(n)},$$

for some constant $c' > 0$. (5)

Finally, combining the results of the above three capacity constraints, equations (3), (4), and (5), and using the assumption that $d(n) \leq s(n)$, i.e., $d(n)/s(n) \leq 1$, we obtain the following upper bound:

$$\begin{aligned} \lambda(n, s(n), d(n)) &= O\left(\min\left(1, \frac{d(n)}{s(n)}, \frac{\sqrt{n/\log n}}{s(n)}\right)\right) \\ &= O\left(\min\left(\frac{d(n)}{s(n)}, \frac{\sqrt{n/\log n}}{s(n)}\right)\right) \end{aligned} \quad (6)$$

5 Constructive Lower Bound on Data Gathering Capacity in Random Networks

In this section we first present a node transmission scheme and routing scheme and derive a constructive lower bound of the data gathering capacity. We then examine the scaling behaviors of the lower bound in different scaling regimes of the number of source and destination nodes. Finally, we investigate the gap between the lower and upper bounds and propose approaches to improve the constructive lower bound and hence narrow the gap.

5.1 Node scheduling scheme

We now construct a node scheduling scheme and show an important property of the schedule.

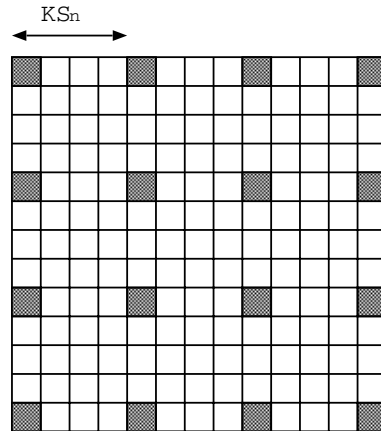


Figure 2. Equivalence classes of squarelets. The unit area square is divided into squarelets of side length $s_n = \sqrt{3 \log n/n}$. Squarelets Ks_n distance apart form an equivalence class. The shaded squarelets form one such equivalence class.

- We divide the unit area square into squarelets of side length $s_n = \sqrt{3 \log n/n}$ as in [10]. It has been proved that all of the squarelet are occupied by at least one node with high probability.
- We then choose the transmission range of each node to be $r = \sqrt{5}s_n$, the maximum distance between two arbitrary points in neighboring squarelets. Therefore, a node can communicate with any node in the neighboring squarelet.
- We further divide the squarelets into K^2 equivalence classes, where K is a positive integer and we require $K \geq 2 + (1 + \Delta)\sqrt{5}$. Each equivalence class is comprised of the squarelets that are exactly K squarelets apart horizontally and vertically, as illustrated in Figure 2.

This construction readily lends itself to a TDMA scheme with the following property.

Lemma 2 *If $K \geq 2 + (1 + \Delta)\sqrt{5}$, there exists a TDMA scheme such that one node per squarelet can transmit to a receiver in a neighboring squarelet at one bit/sec once every constant number (K^2) of time slots.*

Proof. We divide time into slots and let the equivalence classes take turn to transmit. In each turn, one node in each squarelet of the scheduled equivalence class transmits to a receiver in a neighboring squarelet while all other nodes remain silent. Since adjacent squarelets in an equivalence class are Ks_n distance apart from each other, the minimum distance between a receiver and other simultaneous transmitters is $(K-2)s_n > (1+\Delta)\sqrt{5}s_n$, greater than the size of the guarding zone that ensures successful reception at the receiver. Therefore, the simultaneous transmissions in each squarelet of the scheduled equivalence class are all successfully received. Since there are K^2 equivalence classes, a node in each squarelet can successfully transmit to a receiver in a neighboring squarelet once every K^2 time slots. \square

5.2 Routing algorithm and capacity

In the above construction, the unit area square is divided into squarelets of side length $s_n = \sqrt{3 \log n / n}$. There are $\sqrt{n/3 \log n}$ rows and $\sqrt{n/3 \log n}$ columns of squarelets. We index each squarelet by its row and column numbers. Packets from a source node to the destination node are routed in a squarelet-to-squarelet fashion along a Manhattan path. Specifically, packets from a source node in squarelet (i,j) to a destination node in squarelet (u,v) are routed in two phases. (1) from squarelet (i,j) horizontally to squarelet (i,v), and (2) from squarelet (i,v) vertically to squarelet (u,v).

From the construction, the size of the squarelet is large enough so that all squarelets are occupied by nodes whp, and the transmission range of each node is large enough to reach any node in the neighboring squarelets. Therefore, packets can be relayed along the rows and columns continuously throughout the network.

Phase 1: Horizontal relay from source node. In the worst case, a squarelet needs to relay horizontal packets for all of the source nodes that are located on the same row. Since $s(n)$ source nodes are randomly and uniformly distributed in $\sqrt{n/3 \log n}$ rows, the maximum number of sources nodes whose packets go through a squarelet horizontally is $M(s(n), \sqrt{n/3 \log n})$.

From Lemma 2, nodes in a squarelet can relay packets to neighboring squarelets at a rate of 1 bits/sec once every constant number of time slots. Therefore, the capacity of the horizontal relay per source node is $\Omega(1/M(s(n), \sqrt{n/3 \log n}))$ bits/sec.

Phase 2: Vertical relay to destination node. The maximum number of source nodes for a destination node is $M(s(n), d(n))$. Note that if the source nodes are randomly and evenly mapped to the destination nodes, each destination node will be the receiver of exactly $s(n)/d(n)$ source

nodes (for simplicity, assume that $s(n)$ is exactly divisible by $d(n)$).

As in the horizontal relay phase, the maximum number of destination nodes located in each column is $M(d(n), \sqrt{n/3 \log n})$. Therefore, the capacity of the vertical relay per source node is

$$\Omega\left(\frac{1}{M(s(n), d(n))M(d(n), \sqrt{n/3 \log n})}\right) \text{ bits/sec.}$$

Combining the capacity analysis of phase 1 and phase 2, the final constructive lower bound of the data gathering capacity is constrained by the lesser of the two phases. Ignoring the constant factor, we have

$$\lambda(n, s(n), d(n)) = \Omega\left(\min\left(\frac{1}{M(s(n), \sqrt{n/\log n})}, \frac{1}{M(s(n), d(n))M(d(n), \sqrt{n/\log n})}\right)\right)$$

This completes the proof of Theorem 2.

5.3 Scaling behavior of the lower bound

The scaling behaviors of the above capacity lower bound depend on the scaling regimes of $s(n)$ and $d(n)$. Assume $s(n) = \Theta(n^\alpha)$ and $d(n) = \Theta(n^\beta)$ with $0 \leq \beta \leq \alpha \leq 1$. In the following we derive the scaling behavior of the lower bound in different scaling regimes of $s(n)$ and $d(n)$.

- $\alpha \geq \beta > 1/2$.

Based on Lemma 1, the maximum number of source nodes per row in the squarelets structure is

$$M(s, \sqrt{n/\log n}) = s/\sqrt{n/\log n}.$$

The maximum number of source nodes per destination node is

$$M(s, d) = \begin{cases} \frac{1}{n^{\alpha-\beta}} & \alpha > \beta, \\ \frac{\log n}{\log \log n} & \alpha = \beta. \end{cases}$$

The maximum number of destination nodes per column in the squarelets structure is

$$M(d, \sqrt{n/\log n}) = d/\sqrt{n/\log n}.$$

Therefore, the constructive lower bound on the capacity is

$$\lambda(n, s(n), d(n)) = \begin{cases} \Omega\left(\frac{1}{\sqrt{n^{2\alpha-1} \log n}}\right) & \alpha > \beta, \\ \Omega\left(\frac{1}{\sqrt{n^{2\alpha-1} \log n}} \frac{\log \log n}{\log n}\right) & \alpha = \beta. \end{cases}$$

Note that when $\alpha > \beta$, the lower bound is of the same order as the upper bound obtained from equation (1), i.e., the combination of the aforementioned scheduling and routing schemes achieves order-optimal data transport capacity. However, when $\alpha = \beta$, there is a degradation factor of $\log \log n / \log n$. Due to the random mapping between the source and destination nodes, a destination node may serve as many as $\log n / \log \log n$ source nodes, resulting in the capacity degradation. In the special case of $\alpha = \beta = c$ where $0 < c < 1$ is a constant, i.e., $s(n) = d(n) = \Theta(n)$, we have

$$\lambda(n, s(n), d(n)) = \Omega\left(\frac{1}{\sqrt{n \log n}} \frac{\log \log n}{\log n}\right).$$

If a load balancing scheme is used such that there is a one-to-one random mapping between source and destination nodes, i.e., each destination node is the destination of exactly one source node. The capacity scales as $\Theta\left(\frac{1}{\sqrt{n \log n}}\right)$. This recovers the result in the Gupta and Kat work [7] and the Guarani and Viswanath work [10], where the one-to-one random mapping between source and destination nodes is assumed.

- $\alpha > 1/2 > \beta$

In this case we have

$$\begin{aligned} M(s, \sqrt{n/\log n}) &= s/\sqrt{n/\log n} \\ &= \sqrt{n^{2\alpha-1} \log n}, \end{aligned}$$

$$M(s, d) = s/d = n^{\alpha-\beta},$$

and

$$\begin{aligned} M(d, \sqrt{n/\log n}) &= O\left(\frac{\log \sqrt{n/\log n}}{\log \sqrt{n^{1-2\beta}/\log n}}\right) \\ &= O(1). \end{aligned}$$

Therefore, the lower bound on the capacity is

$$\begin{aligned} \lambda(n, s(n), d(n)) &= \Omega\left(\min\left(\frac{1}{\sqrt{n^{2\alpha-1} \log n}}, \frac{1}{n^{\alpha-\beta}}\right)\right) \\ &= \Omega\left(\frac{1}{n^{\alpha-\beta}}\right) \end{aligned}$$

- $1/2 > \alpha > \beta$

In this case we have

$$M(s, \sqrt{n/\log n}) = O(1),$$

$$M(s, d) = s/d = n^{\alpha-\beta},$$

and

$$M(d, \sqrt{n/\log n}) = O(1).$$

Therefore, the lower bound on the capacity is

$$\lambda(n, s(n), d(n)) = \Omega\left(\frac{1}{n^{\alpha-\beta}}\right)$$

We have shown that the constructive capacity lower bound for the case $\alpha > \beta > 1/2$ is order-optimal (i.e., on the same order as the upper bound). It is easy to verify that the same holds for the cases $\alpha > 1/2 > \beta$ and $1/2 > \alpha > \beta$. The scaling laws of the capacity in the different regimes of the source and destination nodes are summarized in Figure 3. Here we only consider the case where $\alpha \geq \beta$. As mentioned in Section 2, the case where $\alpha \leq \beta$ corresponds to the scenario that a smaller number of source nodes send independent information to a larger number of destination nodes, and its capacity is symmetric to the case where $\alpha \geq \beta$.

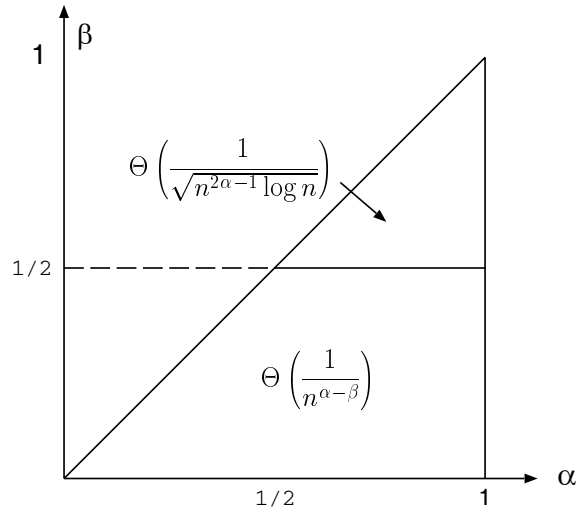


Figure 3. Data gathering capacity lower bound in different scaling regimes of the source and destination nodes.

5.4 Gap between the lower and upper bounds

Above we briefly touched upon the issue that there is a gap between the constructive lower bound of the capacity and the corresponding upper bound in certain scaling regimes of the number of source and destination nodes. Recall that there is a degradation factor of $\log n / \log \log n$ in capacity when $\alpha = \beta > 1/2$ compared to the case when $\alpha > \beta > 1/2$. In the following we further examine the gap between the lower and upper bounds. Based on the analysis, we suggest approaches that can improve the constructive lower bound and hence narrow the gap.

Comparing the analytical expressions of the upper and lower bounds given in equations (1) and (2), it can be observed that a gap between them is possible due to the random locations of the source and destination nodes, as well as the random mapping between the nodes.

First, as the source and destination nodes are randomly selected in the network, these nodes could be geographically concentrated in a small area and need to share the network transport capacity in the area. For the above scheduling and routing schemes, the lower bound is affected by the maximum number of source and destination nodes located in each row or column in the squarelets structure. Note that this effect is not limited to the specific routing scheme (Manhattan routing) used in the construction, but is also present in other deterministic routing schemes.

Second, due to the random mapping between the source and destination nodes, there could be traffic concentration on some of the destination nodes. To this end, the lower bound is affected by the maximum number of source nodes that a destination node receives data from. Based on the maximum occupancy results, the maximum number of source node to a destination node is on the same order of the average number when $s(n) = \Omega(d(n) \log d(n))$, but starts to deviate from the average when $s(n) = O(d(n) \log d(n))$.

The above analysis on the possible gap between the lower and upper bounds readily lends itself to methods of improving the constructive lower bound of the data gathering capacity.

- *Geographic distribution of source and destination nodes.* It is desirable to locate source and destination nodes geographically evenly in the network to avoid node concentration. For example, in a wireless ad hoc or sensor network, it helps to place gateway nodes or sink nodes evenly in the network.

In the construction, the unit area square network is divided into squarelets of side length $\sqrt{3 \log n/n}$. There are $\sqrt{n/3 \log n}$ rows and columns of such squarelets, and a node in each squarelet can successfully transmit to its neighboring squarelet once very constant number of time slots. According to the maximum occupancy results, if the number of source (destination) nodes is asymptotically greater than \sqrt{n} , the maximum number of source (destination) nodes located in a row (column) is on the same order of the average value. In this case there is no significant benefit to relocate the source or destination nodes. On the other hand, if the number of source (destination) nodes is asymptotically smaller than \sqrt{n} , the maximum number of source (destination) nodes per row (column) starts to deviate from the average value. In this case it is beneficial to locate source or destination nodes evenly in the network.

- *Mapping between source and destination nodes.* It is

desirable to map source nodes evenly to destination nodes to avoid traffic concentration. Ideally, we want to have a load balancing scheme that maps the same number of source nodes to each destination node, i.e., each destination node receives data from $s(n)/d(n)$ source nodes. Without load balancing, the maximum number of source nodes that is mapped to a destination node is $M(s(n), d(n))$. Based on the maximum occupancy results, when $s(n) \gg d(n) \log d(n)$, the maximum number of source nodes for the same destination node is on the same order of the average value, $s(n)/d(n)$. The benefit of load balancing is not significant. However, when $s(n) \ll d(n) \log d(n)$, the maximum number of source nodes for the same destination node becomes much larger than the average number. For example, when $s(n) = d(n) = n^\alpha$ for $0 < \alpha \leq 1$, the maximum number of source nodes for the same destination node is $\log n / \log \log n$, much larger than the average value $s(n)/d(n) = 1$. The load balancing at the destination nodes now becomes beneficial.

6 Conclusion

In this paper we study the data gathering capacity of large scale multihop wireless networks. The data gathering communication paradigm represents the network scenarios where a subset of nodes send data to some designated destination nodes, while other nodes serve as relays. This some-to-some communication paradigm is prevalent in many wireless network applications. We derive the upper bound and constructive lower bound of the capacity, and analyze their scaling behaviors in different scaling regimes of the number of source and destination nodes. We find that the capacity is constrained by different factors and exhibits distinct scaling behaviors in the different scaling regimes of $d(n)$. The characterizations of the capacity allows us to identify the bottleneck of the data gathering capacity for given network scenarios and choose proper approaches to improve the capacity. For example, when the capacity is constrained by the number of available destination nodes, it is beneficial to add extra destination nodes. On the other hand, if the capacity is constrained by the network transport capacity to relay the packets, adding more destination nodes will not result in substantial capacity improvement. For future work we plan to study the data gathering capacity when there is some infrastructure (well connected base stations) in the network.

References

- [1] S. Aeron and S. Venkatesh. Scaling laws and operation of wireless ad-hoc and sensor networks. In *IEEE Workshop on Statistical Signal Processing*, 2003.
- [2] H. Bolcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj. Capacity scaling laws in mimo relay networks. *IEEE Transactions on Wireless Communications*, 5(6), June 2006.
- [3] H. E. Gamal. On the scaling laws of dense wireless sensor networks: The data gathering channel. *IEEE Trans. Inform. Theory*, 51(3):1229–1234, March 2005.
- [4] M. Gastpar and M. Vetterli. On the capacity of wireless networks: the relay case. In *Proc. IEEE Infocom*, 2002.
- [5] P. K. Gopala and H. E. Gamal. On the scaling laws of multimodal wireless sensor networks. In *Proc. IEEE Infocom*, 2004.
- [6] M. Grossglauser and D. N. C. Tse. Mobility increases the capacity of ad-hoc wireless networks. In *Proc. IEEE Infocom*, 2001.
- [7] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2), Mar 2000.
- [8] P. Jacquet and G. Rodolakis. Multicast scaling properties in massively dense ad hoc networks. In *Proceedings of the 11th International Conference on Parallel and Distributed Systems - Workshops (ICPADS'05)*, 2005.
- [9] A. Keshavarz-Haddad, V. Ribeiro, and R. Riedi. Broadcast capacity in multihop wireless networks. In *Proceedings of ACM Mobicom*, 2006.
- [10] S. R. Kulkarni and P. Viswanath. A deterministic approach to throughput scaling in wireless networks. *IEEE Trans. on Information Theory*, 2004.
- [11] X.-Y. Li, S.-J. Tang, and O. Frieder. Multicast capacity for large scale wireless ad hoc networks. In *Proceedings of ACM MobiCom*, 2007.
- [12] B. Liu, Z. Liu, and D. Towsley. On the capacity of hybrid wireless networks. In *Proc. IEEE Infocom*, 2003.
- [13] D. Marco, E. Duarte-Melo, M. Liu, and D. L. Neuhoff. On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data. In *Proc. International Workshop on Information Processing in Sensor Networks (IPSN)*, 2003.
- [14] B. S. Mergen and M. Gastpar. On the broadcast capacity of wireless networks. In *Proc. of Information Theory and Applications Workshop*, 2007.
- [15] M. Mitzenmacher. *The Power of Two Choices in Randomized Load Balancing*. PhD thesis, University of California at Berkeley, Berkeley - CA, 1996.
- [16] M. Raab and A. Steger. “balls into bins” — A simple and tight analysis. *Lecture Notes in Computer Science*, 1518, 1998.
- [17] S. Ratnasamy, B. Karp, L. Yin, F. Yu, D. Estrin, R. Govindan, and S. Shenkar. Ght: A geographic hash table for data-centric storage. In *First ACM International Workshop on Wireless Sensor Networks and Applications (WSNA)*, 2002.
- [18] S. Shakkottai, X. Liu, and R. Srikant. The multicast capacity of large multihop wireless networks. In *Proc. of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2007.
- [19] R. Zheng. Asymptotic bounds of information dissemination in power-constrained wireless networks. *IEEE Transactions on Wireless Communications*, 6(12), 2007.