This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Speed Improves Delay-Capacity Trade-Off in MotionCast

Xinbing Wang, *Member*, *IEEE*, Yuanzhe Bei, *Student Member*, *IEEE*,
Qiuyu Peng, *Student Member*, *IEEE*, and Luoyi Fu, *Student Member*, *IEEE*

Abstract—In this paper, we study a unified mobility model for mobile multicast (MotionCast) with n nodes, and k destinations for each multicast session. This model considers nodes which can either serve in a local region or move around globally, with a restricted speed R. In other words, there are two particular forms: Local-based Speed-Restricted Model (LSRM) and Global-based Speed-Restricted Model (GSRM). We find that there is a special turning point when mobility speed varies from zero to the scale of network. For LSRM, as R increases, the delay-capacity trade-off ratio decreases iff R is greater than the turning point $\Theta(\sqrt{\frac{1}{k}})$; For GSRM, as R increases, the trade-off ratio decreases iff R is smaller than the turning point, where the turning point is located at $\Theta(\frac{k^{0.25}}{\sqrt{n}})$ when $k = o(n^{\frac{2}{3}})$, and at $\Theta(\frac{k}{n})$ when $k = \omega(n^{\frac{2}{3}})$. As k increases from 1 to $n-1$, the region that mobility can improve delay-capacity trade-off is enlarged. When $R = \Theta(1)$, the optimal delay-capacity trade-off ratio is achieved. This paper presents a general approach to study the performance of wireless networks under more flexible mobility models.

Index Terms—Ad hoc network, mobility, capacity, delay.

1 Introduction

Delay and capacity trade-off is a fundamental problem in mobile ad hoc networks (MANETs). The notion of capacity is first introduced by Gupta and Kumar [1]. They found optimal static unicast throughput is $\Theta(\frac{1}{\sqrt{n}})$, while

$$\Theta\left(\frac{1}{\sqrt{n\log n}}\right)$$

for random network. In other words, capacity for each node declines when the number of nodes increases. Later, Grossglauser and Tse demonstrated that the per-node capacity can reach $\Theta(1)$ by introducing mobility to the network in [2]. However, such significant gain is achieved at the cost of very large delay. Since both capacity and delay are important factors that determine the performance of a wireless network, researchers paid great effort in understanding the relationship between them in MANETs, namely the delay-capacity trade-off.

Delay-capacity trade-off was first studied by Bansal and Liu [3]. They considered a mobile network with stationary sources and destinations, and proposed a geographic routing scheme to approach the optimal capacity. Later studies, such as Perevalov and Blum [4], studied network capacity in a delay-limited mobile ad hoc network. Recent studies, for example, Neely and Modiano [5], presented a strategy utilizing redundant packets transmissions to reduce delay at the sacrifice of capacity. They established the following necessary trade-off: $delay/capacity \geq O(n)$, and developed schemes that can achieve $\Theta(1)$, $\Theta(\frac{1}{\sqrt{n}})$, and $\Theta(\frac{1}{n\log n})$ per-node capacity, when the average delay is $\Theta(n)$, $\Theta(\sqrt{n})$, and $\Theta(\log n)$, respectively. Toumpis and Goldsmith constructed a better scheme in [6]. Afterwards, Lin and Shroff [7] searched the optimal delay-capacity trade-off and identified the limiting factors of the existing scheduling schemes in MANETs.

One type of work within the delay-capacity trade-off is searching the impact of different mobility models. I.i.d mobility model is the most common and simplest one, thus it has been adopted in many papers, such as [5], [6], [7]. Under this model, each node can move to everywhere in the network area, with uniform probability, and independent of each other. However, in real cases, the node cannot move so fast as it does in i.i.d and i.i.d pattern ignores the mutual dependency and time dependency. So people began to study more realistic mobility models. These include random waypoint mobility model [8], [18]; Brownian mobility model [8], [17], [14]; Random walk mobility model [14], [19]; and Markovian mobility model [5], [15]. In [20], [21], they also study the impact of correlated mobility in MANETs.

When many researchers designed various mobility models to improve the performance in wireless ad hoc networks, people started to consider if there exists a global result that is applicable for all kinds of mobility. Sharma et al. [16] raised a global perspective on delay-capacity trade-off in MANETs. In their paper, a notion called *critical delay* was introduced to explain the nature of node mobility. The result provided a more general understanding of delay-capacity relationship in MANETs.

Another type of work is considering different traffic patterns. Being a more general scheme compared to unicast, multicast scheme and its delay-capacity trade-offs have been studied by many researchers. Li et al. [9] studied the capacity of a static random wireless ad hoc network for multicast. Under protocol interference model, they showed that the per-node multicast capacity is $\Theta(\frac{1}{\sqrt{n\log n}\sqrt{k}})$ when $k = O(\frac{n}{\log n})$,

• *The authors are with the Department of Electronics Engineering, Shanghai Jiao Tong University, No. 800, Dongchuan Road, Shanghai.*
E-mail: xwang8@sjtu.edu.cn, {qiuyupeng, yiluofu}@gmail.com.

Manuscript received 4 Jan. 2010; revised 8 May 2010; accepted 13 May 2010; published online 16 June 2010.
Recommended for acceptance by P. Santi.
For information on obtaining reprints of this article, please send e-mail to: tpds@computer.org, and reference IEEECS Log Number TPDS-2010-01-0019.
Digital Object Identifier no. 10.1109/TPDS.2010.126.

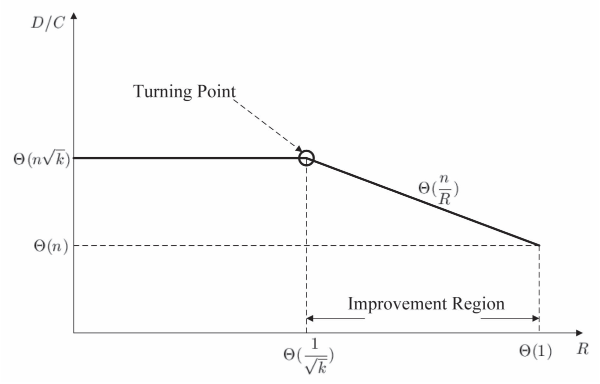


Fig. 1. Curve of the relationship between D/C trade-off and R in LSRM.

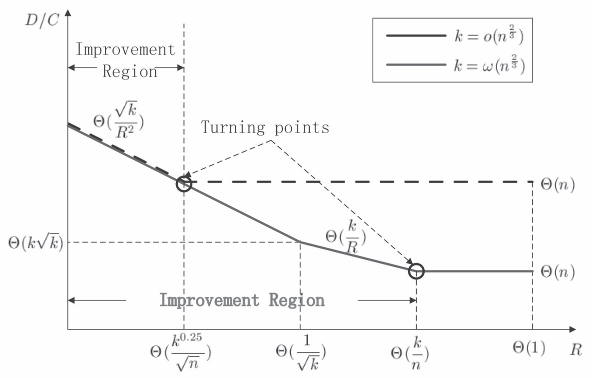


Fig. 2. Curves of the relationship between D/C trade-off and R in GSRM. The upper line and Improvement Region labeled on the top are for $k = o(n^{\frac{2}{3}})$. The lower line and Improvement Region labeled on the bottom are for $k = \omega(n^{\frac{2}{3}})$.

while the per-node multicast capacity is $\Theta(\frac{1}{n})$ when $k = \Omega(\frac{n}{\log n})$. In other studies such as [12], Jacquet and Rodolakis considered multicast capacity by accounting the ratio of the total number of hops for multicast and the average number of hops for unicast, and Shakkottai et al. proposed a comb-based architecture for multicast routing which achieved the upper bound of capacity in [13]. Keshavarz-Haddad et al. studied capacity in broadcast network in [11], which is a special case of multicast. In [23], [24], they study unicast and multicast traffics under Gaussian Channel, respectively. The results obtained can be utilized to better predict the achievable capacity of multicast in a more practical scenario. Very recent study, Hu et al. showed that by two-hop relay strategy, the expectation of delay is no less than $\Omega(\sqrt{n \log k})$ when mobility is introduced and the delay of $O(\sqrt{n \log k})$ is attainable in a proposed scheme with the per-node capacity of

$$\Omega\left(\frac{1}{k\sqrt{n \log k}}\right)$$

in [10]. These results are based on certain mobility models, and therefore they cannot evaluate the network performance when mobility models are in intermediate cases.

In this paper, we present a study considering delay-capacity trade-off for mobile multicast (MotionCast) under general Speed-Restricted mobility models, and focus on how the restricted speed influences delay-capacity trade-off ratio in multicast. We introduce two forms of Speed-Restricted mobility models, under which nodes serve locally or move globally. And we call them Local-based Speed-Restricted Model (LSRM) and Global-based Speed-Restricted Model (GSRM), respectively. We note that in real situations, the restricted mobility speed can also be treated as the restricted range R, in which nodes can move within a single time slot.

Our observation shows that if mobility speed R is located in a specified region, the delay-capacity trade-off can be improved by the increasing R, otherwise the delay-capacity trade-off keeps stable. The relationship between D/C[1] trade-off ratio and R are different under LSRM and GSRM. Moreover, we find that there is a certain turning point, which depends on k, and the increasing R can improve D/C trade-off only when R is greater or less than this turning point.

Our main contributions are summarized as follows:

- Under LSRM, we find that there is a turning point at $R = \Theta(\sqrt{\frac{1}{k}})$ as shown in Fig. 1. When R is smaller than this turning point, the delay-capacity trade-off keeps unchanged; When R is larger than that, such trade-off is improved as R increases.
- Under GSRM, our result shows that the delay-capacity trade-off is divided into two cases depending on k, each of which has a turning point as well. Specifically, 1) for $k = o(n^{\frac{2}{3}})$, the turning point is located at $R = \Theta(\frac{k^{0.25}}{\sqrt{n}})$ as in Fig. 2. Before this turning point, the delay-capacity trade-off is improved as R increases. After that point, such trade-off remains the same; 2) For $k = \omega(n^{\frac{2}{3}})$, the turning point is located at $R = \Theta(\frac{k}{n})$. The phenomenon brought by this turning point is the same as that in the first case except for another point at $R = \Theta(\sqrt{\frac{1}{k}})$, which separates different descending rate on either side.
- For both LSRM and GSRM, speed R can improve D/C trade-off in a wider region as k increases, shown as the "Improvement Region" in Figs. 1 and 2. Particularly, in unicast, i.e., $k = 1$, the D/C trade-offs always keep stable and they are not influenced by mobility speed R. In contrast, in broadcast,[2] i.e., $k = n - 1$, the D/C trade-offs always decrease with the increase of mobility speed R.

The rest of paper is organized as follows: In Section 2, we describe the network model. In Section 3, we define redundancy and find out how it influences the throughput capacity. In Section 4 and 5, we study delay-capacity trade-off under LSRM and GSRM, respectively. In Section 6, we discuss our findings. Finally, we make conclusion and establish future work in Section 7.

1. D/C stands for Delay/Capacity.

2. Because the turning point now is $\Theta(1/\sqrt{n})$, which is just the origin of the coordinates.

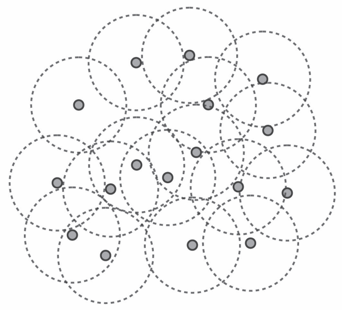


Fig. 3. LSRM.

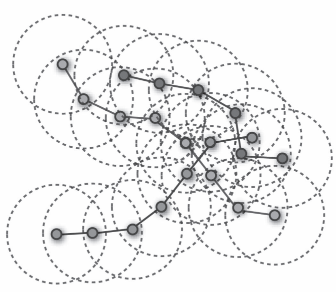


Fig. 4. GSRM.

2 Network Model

2.1 Transmission Model

In our paper, we assume a similar transmission protocol model raised in [1], it is derivative from a widely used physical interference model as Shannmon Capacity:

$$C = W \log_2(1 + SINR),$$

where SINR is the signal to interference and noise ratio. And our proposed model is identical to this famous model, which is also proved in [1]. The whole network is placed in a 1×1 sized square area, with totally n nodes. All of these nodes are uniformly and randomly deployed at the very beginning. We use protocol interference model, which means node v_i can transmit data to node v_j, iff their mutual distance satisfies $d(v_i, v_j) < r$. And any other node $v_k \neq v_i, v_j$ that within v_j's transmission range, namely $d(v_k, v_j) < r$, cannot transmit simultaneously. We assume that each node transmits exactly one packet to exactly one receiver in a single time slot, if this node is permitted to transmit and transmission errors can be avoided if the transmission protocol is satisfied. The sending node cannot transmit a fraction of one packet. We also do not consider network coding strategy so that a sender cannot transmit data to various receivers in a same time slot.

We assume that the buffer size is infinity in all wireless nodes. So each node can keep the packets it has received for quite a long time. The packets are held in the buffer until they are no longer needed to be transmitted to any other nodes. This assumption indicates that each node will never receive a same packet more than once.

Another important assumption is that the network uses critical transmission range. Many studies, such as [1], [2], [5], [9], [10], used critical transmission range as real transmission range. Note that the smallest transmission range may not always result in optimal performance. For example, Lin and Shroff discussed larger transmission range can sometimes reduce the network delay-capacity trade-off ratio in [7]. Note that the reason for us to utilize this model is that the problems raised in this paper is to study asymptotic capacity and delay trade-offs, thus we can assume a constant radio transmission range. Because even if the power of radio is not identical, we can utilize a constant c_0 and c_1 to upper bound and lower bound the power, and this can merely affect our results up to a constant factor, which would not affect the order of our results.

2.2 LSRM

We propose a Speed-Restricted Model (SRM) which has two particular forms. In both of them, the speed of nodes are restricted by R. Suppose nodes are distributed uniformly and randomly at the beginning. For LSRM, they can only move within their own circle area centered at their initial position at each time slot. We set the radius of these circles as R so that the mobility speed of nodes is restricted by R (See Fig. 3).

Assume each node can move i.i.d within its own circle. So it can be treated as a small scaled i.i.d mobility pattern. Notice that when the radius reaches $\Theta(1)$, the corresponding mobility model is i.i.d, and when the radius is restricted as $O(\sqrt{\log n/n})$, the network is equivalently stationary.[3] Thus, we can unify these two mobility patterns by LSRM.

The LSRM requires all the nodes must stay in a certain area, which indicates that the moving areas of the source and the destination may not overlap. Therefore, relays are required in this pattern. Intuitively, the whole route is covered by a chain of circles, with every two adjacent circles overlap.

2.3 GSRM

In LSRM, the centers of all circles where nodes move around are fixed after the nodes have been initially deployed. However, in GSRM, we also let the centers of these circles move. They are the positions of the nodes in the previous time slot. Since speed is restricted by R, the radius of these circles are also R. Nodes can move to any point of the current circles in uniform probability at each time slot.

Obviously, all the nodes can reach the whole network region in this model, but the distance of each hop is restricted by R. Notice that when R is zero, the network is stationary; when $R = \Theta(\sqrt{1/n})$, the mobility model can be treated as Markov Model;[4] when $R = \Theta(1)$, the corresponding mobility pattern comes to i.i.d again. Fig. 4 shows the movement of three different nodes.

3 Fundamental Theorems of Capacity and Delay

In this section, we provide some fundamental theorems that are essential throughout the whole paper. First, we find out the relationship among redundancy, critical transmission range, and capacity. Therefore, we can divide each problem

3. Connectivity can be guaranteed when the transmission range $r = \Omega(\sqrt{\frac{\log n}{n}})$, which is proved in original work. Therefore, when $R = O(\sqrt{logn/n})$, the node can be regarded as stationary because each node is within the transmission range of a certain chosen node w.h.p.

4. The reason is provided in [17].

about capacity into two subproblems, one is about redundancy and the other is about critical transmission range. Besides, we prove some frequently used lemmas and theorems in this section.

3.1 Definition of Redundancy

Redundancy is an important concept in wireless network. The same packet maybe transmitted more than once. This could be three reasons: multihop, data-copy, and multicast scheme. We define redundancy as the average number of transmissions for each packet in the network.

Multihop is commonly used in stationary network. When the source and the destination are so far separated that such distance exceeds the source's transmission range, relays are necessary. A group of relays are chosen from the network and then form a chain route for the packet traveling from the source to the destination. Actually, in the network if nodes cannot traverse to every position in the network area, multihop is necessary. When the chain is composed of one source, one destination, and $m-1$ relays, the packet will be transmitted for m times, and the redundancy is therefore m.

Data-copy is another kind of redundancy, which is widely used in mobility network, such as i.i.d network. In MANET, if the source transmits data to the destination by ceaseless moving until their separation distance is shorter than transmission range, it takes very long time to finish a session when the transmission range is small. To reduce the delay, we choose a group of nodes as relays. The source hands out its packet to all these relays. Since there is much more chance for the destination to meet many relays than to meet a single source, the overall delay decreases. Suppose each source hands out its packet to $m-1$ relays, and the packet is transmitted for m times, including $m-1$ transmissions for handing out and once for one of the relays transmits the packet to the destination.

Different from unicast, redundancy can also happen in multicast networks. The multicast network requires each packet has k different destinations. Obviously, there are at least k transmissions in each session even no multihop and data-copy are used. That is why the capacity of multicast network is lower than that of unicast network.

Consider any ad hoc network, a session starts with packet P. Let v_s denote the source that emits P. Let $(v_i, v_j)_t$ denote the transmission pair where node v_i transmits packet P to node v_j at time slot t. Consider the set of all transmission pairs that transmit packet P: $E_P^* = \{(v_i, v_j)_t\}$. And then, $|E_P^*|$ is the redundancy for packet P. However, if we consider different sessions, $(v_i, v_j)_{t_1}$ and $(v_i, v_j)_{t_2}$ are different transmission pairs when $t_1 \neq t_2$.

Consider the set of edge $E_P = \{\overline{v_i v_j} | \exists t > 0, (v_i, v_j)_t \in E_P^*\}$ and the set of vertex V_P contains all the points that perform as senders or receivers in E_P^*. The graph $G_P(V_P, E_P)$ therefore presents the transmission flow for packet P. Each packet has its own transmission flow, and the number of redundancy is the number of nodes in graph G_P. G_P has following characteristics (proved in Technical Report [7]):

- $|E_P^*| = |E_P|$. We only need to prove that the pairs $(v_i, v_j)_{t_1}$ and $(v_i, v_j)_{t_2}$ cannot both exist in $|E_P^*|$. In fact, if we assume $t_1 < t_2$, node v_j will have already holden packet P at time slot t_2. So transmission pair $(v_i, v_j)_{t_2}$ is meaningless.
- G_P is connected. Because only v_s produce the packet P. If G_P is not connected, the nodes that are not connected to v_s will have no opportunity to get packet P.
- All nodes except v_s in G_P has input degree one. Since buffer size is infinity in each wireless node, once a node get packet P, it will hold it until the entire session for P accomplishes, and therefore no longer need to receive P from another sender.
- G_P is a tree. We only need to prove there are no rings in G_P. If there exists one, $(v_1 \to v_2 \to \cdots \to v_k \to v_1)$. Suppose v_1 is the first one that gets the packet P, then v_k must gets packet P later than v_1, so transmission pair (v_k, v_1) is not necessary.

To sum up, although there are various situations that may require redundancy in transmissions, the essence of redundancy is the same: the number of direct transmissions with same packets, or the size of transmission flow tree for each packet. In the next part, we will figure out the relationship between redundancy and network capacity.

3.2 Redundancy Influences Network Capacity

In the previous part, we discuss the format of redundancy in all kinds of wireless ad hoc networks. In this part, we study the relation between redundancy and network capacity.

First, we recall the definition of network capacity. The overall network capacity is the average number of packets that can be transmitted simultaneously in the network. However, in most cases, all transmission pairs serve different packets (i.e., $m = 1$) is impossible. Just as we mentioned in the previous part, redundancy is necessary in many situations.

Theorem 3.1. *The network capacity is inversely proportional to the redundancy. Suppose there are n_s transmission pairs available in each time slot. Then, the whole capacity is $\Theta(\frac{n_s}{m})$ if the redundancy for each packet is m.*

Proof. We consider the network has been working for T time slots. Then, the network provides n_sT peer-to-peer transmissions pairs for us (same pair in different time slots are regarded as different pairs). Notice that every successful source-to-destination transmission needs average m peer-to-peer transmission.[5] Therefore, there are utmost $\frac{n_sT}{m}$ transmissions can be achieved during these T time slots. Then, let T goes to infinity, the per-node capacity is the average number of source-to-destination transmissions in the network, namely $c = \Theta(\frac{n_s}{m})$.

Then, we will try to find a scheduling scheme to achieve this upper bound. There must exist an algorithm that can make full use of n_sT transmission pairs in T time slots. Suppose we can know all these transmission pairs and write them in a data array, which is sorted by their happening time. Then, our scheduling algorithm is designed as follows:

- Step 1: First we sort the sessions according to its arrival time. We place a higher prior to sessions

5. We must allow for m duplicates of the packet according to our scheme and every additional duplicate requires a successful transmission.

beginning earlier than those beginning later. For instance, if two transmission pair $(v_i, v_j) \in Session_m$ and $(v_k, v_l) \in Session_n$, respectively, and $Session_m$ begins earlier than $Session_n$, then $Session_m$ is higher than $Session_n$. And the transmission pair (v_i, v_j) is also higher than (v_k, v_l).

- Step 2: Then, we sort all the transmission pair in the same session G_p, traverse every transmission pair from the root to leaves with width search. Then, we obtain a list of visiting sequence and assume pair (v_i, v_j) is in the m_{th} position in the sequence and (v_k, v_l) in the n_{th} position, if $m < n$, then we place a higher prior to (v_i, v_j) than (v_k, v_l).
- Step 3: In every new time slot, we would search a set of n_s transmission pairs in the data array according to the priorities which also satisfy the transmission protocols. After transmission, we would remove it from the data array until the array is empty. Assume that each session can at most last $T_0 = o(m)$ sessions, then if T is large enough, the achievable capacity is

$$C = \lim_{T \to \infty} \frac{n_s T - n_s T_0}{Tm} = \Theta\left(\frac{n_s}{m}\right).$$

Therefore, under this algorithm, and let T go to infinity, each transmission pair can serve one packet for communication. The network resource is therefore fully used. □

3.3 Critical Transmission Range Influences Network Capacity

In the previous part, we know that the redundancy is inverse proportional to the network capacity, namely $c = \Theta(\frac{n_s}{m})$. In this part, we find out the value of n_s, which is determined by critical transmission range r.

Critical transmission range r is an essential concept in ad hoc wireless network. We define this range as the minimum transmission range that can make the network connected w.h.p. We call a network connected if there is no subset S of the node set V, such that every node $v_i \in S$ and $v_j \in \overline{S}$ are separated larger than r, namely $d(v_i, v_j) > r$.

The critical transmission range varies from $\Theta(\sqrt{\frac{1}{n}})$ to $\Theta(\sqrt{\frac{\log n}{n}})$, which are feasible in i.i.d mobility network [5] and stationary network [1], respectively. The cause of variation between critical transmission ranges is mobility, namely mobility can improve connectivity.

According to the definition of transmission range and protocol interference model, when a sender v_i is sending data to a receiver v_j, all other nodes that within v_j's transmission area cannot send packet at the same time. In other words, there only exists one transmission pair in v_j's transmission area. In Fig. 5, we show that the total number of transmission pairs is determined by the number of nonoverlapping circles with radius r in the whole network area. The following theorem determines the number of nonoverlapping circles.

Theorem 3.2. *We can choose at most $\Theta(\frac{1}{r^2})$ nodes from the network w.h.p so that the circles that are centered at the chosen nodes with radius r do not overlap to each other.*

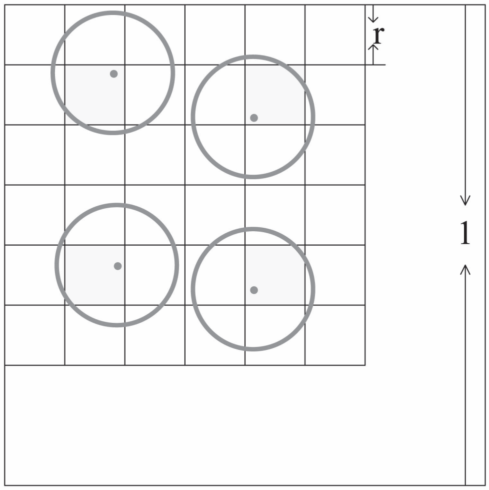


Fig. 5. Nonoverlapping circles.

Proof. We divide the proof into two steps. First, we consider the upper bound. The area for each active circle is πr^2. Therefore, the number of nonoverlapping circles cannot exceed $\frac{1}{\pi r^2}$. The upper bound is $\Theta(\frac{1}{r^2})$.

Then, we come to the lower bound. Assume the network area, a unit square is separated into subsquares with side length $3r$. Then, there are totally $\frac{1}{9r^2}$ subsquares. Then, we divide one of the subsquare into nine squares with side length r. If a node whose central point is located in the central square, it is obvious that this node's entire active area is within the subsquare.

Then, there are totally $\frac{1}{9r^2}$ central squares. We only need to prove that $\Theta(\frac{1}{9r^2})$ of them contain at least one node in it. Since the nodes are deployed randomly and uniformly into the network area, the probability that no nodes inside one of the central square is

$$(1 - r^2)^n < \left(1 - \frac{1}{n}\right)^n \to \frac{1}{e}.$$

Notice that this probability goes to a constant when n goes to infinity. And we use $r > \sqrt{\frac{1}{n}}$ because the radius r need to be larger than average distance between adjacent nodes. For each cell, the probability that the cell contains at least one node is no less than $1 - \frac{1}{e}$. Therefore, there are $\frac{e-1}{9er^2}$ central squares can contain at least one node with high probability. Thus, the lower bound is also $\Theta(\frac{1}{r^2})$. Then, the theorem is proved. □

According to Theorem 3.2, the number of nonoverlapping circles with radius r in the network area is $\Theta(\frac{1}{r^2})$. Then, the number of transmissions that can be maintained simultaneously is also $\Theta(\frac{1}{r^2})$. If there is no redundancy, the overall network capacity is bounded with $\Theta(\frac{1}{r^2})$.

3.4 Relationship among Critical Transmission Range, Redundancy, and Capacity

In the previous two parts, we show that network capacity c and redundancy m have relation $c \propto \frac{1}{m}$, while network capacity c and critical transmission range r have relation $c \propto \frac{1}{r^2}$. When there is no redundancy, namely for each data packet P, there are only one source, one destination, and no relays, the overall capacity is $\Theta(\frac{1}{r^2})$. When the redundancy is m, each packet need m transmission pairs, so the capacity has a factor of m degeneration. The relationship between redundancy m, per-node capacity c, and critical transmission range r can be displayed as follows:

$$c = \Theta\left(\frac{1}{nmr^2}\right), \tag{1}$$

where n denotes the total number of nodes. Equation (1) shows us a general way to determine network capacity. So we only need to determine the redundancy and critical transmission range, which is much simpler. This equation is applicable for many kinds of wireless ad hoc networks, e.g., [1], [5], [9], [10]. We evaluate their results by our formula as follows.

3.4.1 Stationary Unicast Network

In [1], author shows the per-node capacity is $\Theta(\sqrt{\frac{1}{n\log n}})$. The number of hops is $\Theta(\sqrt{\frac{n}{\log n}})$. And the critical transmission range is $\Theta(\sqrt{\frac{\log n}{n}})$. Thus, (1) is satisfied.

3.4.2 Stationary Multicast Network

The stationary multicast network is quite similar as that of the unicast scheme. The only difference is that the transmission route is no longer a chain but a tree. And the redundancy is just the size of the spanning tree. In [9], the author proved that the tree size is on the order of $\Theta(\sqrt{\frac{n}{\log n}} \cdot \sqrt{k})$ when $k = O(\frac{n}{\log n})$. Since the critical transmission range is only based on the deployment, so it is $\Theta(\sqrt{\frac{\log n}{n}})$, as well as the unicast case and (1) is still satisfied.

3.4.3 The i.i.d Network

In [5], the result is that the per-node capacity is $\Theta(1), \Theta(\frac{1}{\sqrt{n}})$, and $\Theta(\frac{1}{n\log n})$ when the delay constraint is on order of $\Theta(n)$, $\Theta(\sqrt{n})$, and $\Theta(\log n)$, respectively. Notice that the critical transmission range of i.i.d ad hoc network is $\Theta(\sqrt{\frac{1}{n}})$. And its redundancy, however, entirely depends on data copies with one-hop and two-hop strategy, while has a factor of $\Theta(\log n)$ in multihop strategy.

The redundancy of one-hop strategy, for example, is $\Theta(1)$ for there are no redundancy at all. Therefore, according to (1), the corresponding per-node capacity is $\Theta(1)$.

When using two-hop strategy, the author shows the number of data copies is $\sqrt{n}$. The redundancy is therefore $\sqrt{n}$ as well. Obviously, (1) still holds.

Finally, consider the multihop strategy. One specialty is that the number of hops is no longer a constant, and therefore should be considered. Notice that the flooding algorithm is used in this model, the source node need to sent its packet to all the other $n-1$ nodes, so the number of data copies is $\Theta(n)$. Then, the overall redundancy is the combination of data copies redundancy and multihop redundancy, namely $\Theta(n\log n)$. This can explain why the per-node capacity is $\Theta(\frac{1}{n\log n})$ for multihop strategy in i.i.d ad hoc network.

3.5 Data-Copy Influences Capacity and Delay in i.i.d Networks

In MANET, we can use data-copy strategy to reduce delay by sacrificing some capacity. Compared with the strategy that only the source can meet the destination, data-copy provides larger probability and therefore requires shorter delay. In this part, we study how data-copy influences capacity and delay in i.i.d network.

For unicast, let v_s denote the source, v_d denote the destination, and $r_1, r_2, \ldots, r_m$ denote m relays. Consider these relays' transmission area $\Psi_{r_1}, \Psi_{r_2}, \ldots, \Psi_{r_m}$, which are all circular with radius r. These m circle areas form a combined area $\Psi = \Psi_{r_1} \cup \Psi_{r_1} \cup \cdots \cup \Psi_{r_m}$. The event that v_d can receive data from one of the relay is equal to the event that v_d is inside Ψ. By determining the area of Ψ, namely $S(\Psi)$, we can know the probability that the destination node can receive packet.

Theorem 3.3. *If $r = \Theta(\sqrt{\frac{1}{n}})$, $m < n$, then $S(\Psi) = \Theta(mr^2)$.*

Proof. First consider the upper bound. Obviously, we have

$$\begin{aligned} S(\Psi) &= S(\Psi_{r_1} \cup \Psi_{r_1} \cup \cdots \cup \Psi_{r_m}) \\ &\leq S(\Psi_{r_1}) \cup S(\Psi_{r_1}) \cup \cdots \cup S(\Psi_{r_m}) \\ &= m\pi r^2 \\ &= \Theta(mr^2). \end{aligned} \tag{2}$$

Then, we consider the lower bound. For $m = \Theta(1)$, we have $S(\Psi) > S(\Psi_{r_1}) = \pi r^2$.

For $m = \omega(1)$, since relay nodes $r_1, r_2, \ldots, r_m$ are uniformly distributed in the network area, we can divide the whole network area into $\sqrt{m} \times \sqrt{m}$ cells. For each cell, the probability that it does not contain any of these m relays is $(1 - \frac{1}{m})^m \to \frac{1}{e}$, which is a constant. According to Law of large numbers, the number of cell that contains relay nodes is on the same scale of the total number of cells, namely $\Theta(m)$. If $m = o(n)$, since $S(\Psi_{r_j}) = \pi r^2 = \Theta(1/n) = o(1/m)$, which indicates that the area of circle is much smaller than the area of single cell. Thus, we can treat the circles that are located in different cells do not overlap. Then, we have $S(\Psi) > \Theta(m) \cdot \pi r^2 = \Theta(mr^2)$; if $m = \Theta(n)$, we have $\sqrt{\frac{1}{m}} = \Theta(r^2)$, then if one cell contains at least one relay node, it will be covered at least $\min(\frac{1}{m}, \frac{\pi r^2}{4})$. Therefore, $S(\Psi) > \Theta(m) \cdot \Theta(r^2) = \Theta(mr^2)$. So the lower bound is proved. To sum up, $S(\Psi) = \Theta(mr^2)$. □

Then, the following theorems show how data-copy strategy influences network capacity and delay.

Theorem 3.4. *Consider a unicast network with per-node capacity c and delay d, under i.i.d mobility model and with no data-copy. If we use m-data-copy ($m = o(n)$), the delay is $\Theta(\frac{d}{m})$, and the per-node capacity is $\Theta(\frac{c}{m})$.*

Proof. According to (1), the network capacity is in the inverse proportion to the redundancy. Therefore, the per-node capacity is $\Theta(\frac{c}{m})$.

Then, we consider delay. The total delay is composed by two steps. In the first step, the source node broadcasts its packet to m relay nodes,[6] this step takes 1 time slot. In the second step, m relays continue to move until one of which meet the destination node. According to Theorem 3.3, the

6. A proper transmission radius can help to achieve this goal because a larger transmission range can result in a larger number of nodes receiving the packet.

area that m relays can transmit data is m times larger than the area that the source node can transmit data, namely $S(\Psi) = \Theta(mr^2) = m \cdot \Theta(r^2)$. Then, the probability for the event that the destination node can get the packet is m times larger after we use m-data-copy strategy, and therefore the delay will decrease by m times. Thus, the delay for second step is $\Theta(\frac{d}{m})$. So the total delay is on the order $\Theta(\frac{d}{m} + 1) = \Theta(\frac{d}{m})$. □

Theorem 3.5. *Consider a k-multicast network with per-node capacity c and delay d, under i.i.d mobility model and with no data-copy. If we use m-data-copy ($m = o(n)$), the delay is $\Theta(\frac{kd}{m+k})$, and the per-node capacity is $\Theta(\frac{kc}{m+k})$.*

Proof. According to (1), the network capacity is in the inverse proportion to the redundancy. Because the redundancy of k-multicast with data-copy scheme is $k + m$ while that without data-copy scheme is k. The per-node capacity is $\Theta(\frac{kc}{m+c})$.

Then, we consider delay. Compare the average delay to meet a new destination with either schemes, assume m' destinations have received the packet. Let d' denote the average delay for unicast without data-copy. For non-data-copy scheme, there are totally $1 + m'$ packet holders and $k - m'$ unsent destinations. Therefore the average delay to meet a new destination is $\Theta(\frac{d'}{(m'+1)(k-m')})$. For data-copy scheme, there are totally $m + m' + 1$ packet holders and $k - m'$ unsent destinations. So the average delay for a new transmission is $\Theta(\frac{d}{(m+m'+1)(k-m')})$. By summing up all the terms as m' vary from 0 to $k - 1$, the Theorem holds. □

4 Capacity and Delay in LSRM

In this section, we study the performance of network under LSRM. We mainly discuss multicast scheme. When k is set as 1, the result can cover that in unicast. Our result can also cover that in stationary unicast network [1], stationary multicast network [9] and i.i.d unicast network [5].

We have the following main results in this section:

- For $R = O(\frac{1}{\sqrt{k}})$, non-data-copy scheme, per-node capacity is
$$\Theta\left(\frac{R}{\sqrt{k}\log(1/R)}\right),$$
and delay is $\Theta(\frac{nR}{\log(1/R)})$.
- For $R = O(\frac{1}{\sqrt{k}})$, data-copy scheme, per-node capacity is
$$\Theta\left(\frac{R}{m\sqrt{k}\log(1/R)}\right),$$
and delay is $\Theta(\frac{nR}{m\log(1/R)})$.
- For $R = \Omega(\frac{1}{\sqrt{k}})$, non-data-copy scheme, per-node capacity is
$$\Theta\left(\frac{1}{k\log(1/R)}\right),$$
and delay is
$$\Theta\left(\frac{n\log(kR^2)}{kR\log(1/R)}\right).$$
- For $R = \Omega(\frac{1}{\sqrt{k}})$, data-copy scheme, per-node capacity is
$$\Theta\left(\frac{1}{(k+m)\log(1/R)}\right),$$
and delay is $\Theta(\frac{n\log(kR^2)}{(k+m)R\log(1/R)})$.
- When $k = 1$, the answer is for unicast scheme.

4.1 Generalized Transmission Range

Generalized transmission range is defined as the largest range that nodes can communicate via mobility. This range consists of radius of moving area and node's transmission range. Since the radius of active area is much larger than transmission range itself, we simply regard the node's transmission range as the radius of active area.

When transmission range R is settled, we have to find out how many nodes can this node communicate directly. Since the network is deployed randomly and uniformly, the average number of nodes inside an active area is $n\pi R^2$. However, we cannot ensure there are $n\pi R^2$ nodes in the area w.h.p for all active areas. Under this circumstance, it is hard to determine how many nodes are available in scheduling algorithm, and the network may not be connected w.h.p when $r = \Theta(\sqrt{1/n})$.

Theorem 4.1. *We can ensure at least $\Theta(\frac{nR^2}{\log(1/R)})$ nodes are placed in a circle with radius R w.h.p.*

Proof. If $q = \Theta(n)$, the lower bound of R is $\Theta(1)$ because the average density of network is $\Theta(\frac{1}{n})$ and then we require $R = \Omega(\sqrt{q/n}) = \Omega(1)$. Therefore, we have $R = \Theta(1)$ because the side length of the whole network area is 1. Obviously, in this case, $q = \Theta(n)$ satisfies the theorem.

If $q = o(n)$, obviously $q < n\pi R^2$. The probability of the event that at least q nodes are in the area is

$$\begin{aligned} P &= \sum_{i=0}^{n-q} C_n^{q+i}(1-\pi R^2)^{n-q-i}(\pi R^2)^{q+i} \\ &= (1-\pi R^2)^{n-q}\sum_{i=0}^{n-q} C_n^{q+i}\tau^i, \end{aligned} \tag{3}$$

where $\tau = \frac{\pi R^2}{1-\pi R^2}$. Use Stirling Formula

$$n! = \Theta(\sqrt{2\pi n}e^{-n}n^n)$$

and get

$$\begin{aligned} \log P &\approx (n-q)\log(1-\pi R^2) + \max\big(\log C_n^{q+i} + i\log\tau\big) \\ &\approx -(n-q)\pi R^2 + \max(n\log n - (q+i)\log(q+i) \\ &\quad - (n-q-i)\log(n-q-i) + i\log\tau). \end{aligned}$$

The maximum term is achieved when $i = n\pi R^2 - q$.

$$\begin{aligned} \log P &\approx -(n-q)\pi R^2 + n\log n - (n\pi R^2)\log(n\pi R^2) \\ &\quad - (n - n\pi R^2)\log(n - n\pi R^2) + (n\pi R^2 - q)\log\tau \\ &\approx -(n-q)\pi R^2 + n\pi R^2(\log n - \log n\pi R^2) \\ &\quad + (n\pi R^2 - q)\log\tau. \end{aligned} \tag{4}$$

Notice that $\log \tau = \log \pi R^2 - \log(1 - \pi R^2) \approx -2\log(1/R)$.

$$\begin{aligned}\log P &\approx -(n-q)\pi R^2 + 2n\pi R^2 \log(1/R) \\ &\quad - 2(n\pi R^2 - q)\log(1/R) \\ &= -(n-q)\pi R^2 + 2q\log(1/R).\end{aligned} \tag{5}$$

Consider boundary condition $\log P = 0$, we get

$$q = \frac{n\pi R^2}{2\log(1/R) + \pi R^2} = \Theta\left(\frac{nR^2}{\log(1/R)}\right). \tag{6}$$

4.2 Critical Transmission Range

In this part, we determine the critical transmission range r under LSRM with radius R. According to Theorem 4.1, the critical number of nodes in a circle area with radius R is $\Theta(\frac{nR^2}{\log(1/R)})$, which is lower than the average $\Theta(nR^2)$, with a degeneration factor $\log(1/R)$. The larger the radius is, the shorter the gap between critical number of nodes and the average number of nodes inside a circular area with radius R.

Suppose a circular area with radius $2R$ contains $c_1 \frac{nR^2}{\log(1/R)}$, where c_1 is a specified constant. Obviously, all the nodes that are outside the circular area with radius $2R$ can never enter its concentric circle with radius R. Therefore, the number of nodes within the inner circle will never exceed that amount.

Theorem 4.2. *The critical transmission range of LSRM is*

$$r = \Theta\left(\sqrt{\frac{\log(1/R)}{n}}\right).$$

Proof. Let $n' = c_1 \frac{nR^2}{\log(1/R)}$, consider the circle that contains n' nodes, which are labeled as $v'_1, v'_2, \ldots, v'_{n'}$. Divide this area into n'/d cells with area $\frac{4\pi r^2}{d}$. In each cell, there are averagely d nodes. Then, the problem is similar to that in the i.i.d model.

Consider the probability that a certain cell is empty, which is $(1 - \frac{d}{n'})^{n'} \to e^{-d}$. And the probability that this cell only contains one node is $\binom{n'}{1}(\frac{d}{n'})(1 - \frac{d}{n'})^{n'-1} \to de^{-d}$. We say a node v'_i is isolated if no other node is within its transmission range. Then, the probability that v'_i is isolated is a constant according to the above two equations. We label it as p, where $0 < p < 1$. Given a long period of time t, the probability that v'_i is isolated is p^t, which goes to zero as t goes to infinity, so the probability that v'_i is isolated from the network permanently is zero, which means v'_i is connected to the network.

Then, we prove that such transmission range is necessary. Suppose we introduce a transmission range

$$r' = o\left(\sqrt{\frac{\log(1/R)}{n}}\right)$$

to the network, then the average number of nodes inside each cell is $o(1)$. This means when the number of nodes n goes to infinity, the average number of nodes in each cell goes to zero. It is easy to get the probability that a certain node v'_i is isolated can approach 1. Therefore, the network is not connected. □

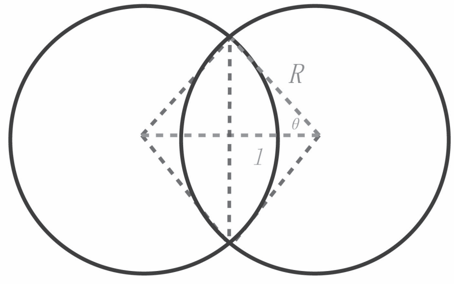


Fig. 6. A transmission pair.

When R is set to be $\Theta(\sqrt{\frac{n}{\log n}})$, r is identical to the critical transmission range of stationary network [1]; when R is set as $\Theta(1)$, r is equal to that in i.i.d mobility network [5].

4.3 Number of Hops

Multihop is necessary in LSRM because nodes have limited active areas. When the source and the destination are separated larger than $2R$, they cannot transmit data directly. Multihop is one type of redundancy, which influences capacity. The number of hops is determined by the average length of each hop and the average distance between source and destination.

Theorem 4.3. *If two nodes mobile uniformly and independently within a same circle of radius R, the delay for the packet sent from one to the other is d (namely, the average time until mutual distance between them is less than r). Or if two nodes move within their own moving areas, the delay for the packet sent from one to the other is d′. If the centers of two circles are separated strictly closer than 2R, we have d′ = Θ(d).*

Proof. As shown in Fig. 6, let l denotes the distance between centers of two circle. The two circles overlap only if $l < 2R$. The overlapping area is determined by R and l. We can get its area

$$\begin{aligned}S_{overlap} &= 2(S_{sec} - S_{tri}) \\ &= 2\left(\theta R^2 - \frac{1}{2}R^2 \sin 2\theta\right) \\ &= R^2(2\theta - \sin 2\theta) \\ &= C_\theta R^2 \\ &= \Theta(R^2),\end{aligned} \tag{7}$$

where $C_\Theta = 2\theta - \sin 2\theta$ is a constant determined by θ, when $l < 2R$. We can see the overlapping area is the same order as the area of circle with radius R. As a result, both the sender and the receiver can enter the overlapping area with probability $\frac{2\theta - \sin 2\theta}{\pi}$, and the probability that both these two nodes enter the overlapping area is $(\frac{2\theta - \sin 2\theta}{\pi})^2$, which is also a constant. According to the law of large numbers, there exists a constant time length $\Delta d = \Theta(1)$, such that with high probability, the sender and the receiver are both inside the overlapping area in at least one time slot in continuous Δd time slots. If d denotes average delay of unicast in circle area with radius R, the delay of unicast in this model does not exceed $d \cdot \Delta d = \Theta(d)$. The order does not change. □

So we only need to limit the hop length l strictly lower than $2R$. The larger l is, the less number of hops we require, and the less probability for the two nodes meet each other. In order to make this probability larger, we can even apply some limitations in network scheduling algorithm. We only allow the pairs whose original separation distance shorter than αR to transmit, where $\alpha < 2$ is a constant parameter. It can be easily proved that the average distance between randomly chosen sender and receiver is $\Theta(1)$. Therefore, the average number of hops is $\Theta(\frac{1}{R})$.

4.4 Non-Data-Copy Scheme

When a new session begins, one source and $k-1$ destinations are selected in the network. Then, the system will build a spanning tree with average branch length $\Theta(R)$. Our task is to determine the total number of branches in this spanning tree.

First, we define the source and the destinations in a session as bifurcate nodes. Consider a spanning tree based on all k bifurcate nodes. We call this spanning tree the original spanning tree. The length of minimum spanning tree $L(k)$ satisfies $E(L(k)) = \Theta(\sqrt{k})$ [22, Lemma 3.3]. Therefore, the average length of each branch is $\Theta(\sqrt{\frac{1}{k}})$.

4.4.1 Capacity

If $R = O(\sqrt{\frac{1}{k}})$, the length of each hop is lower than the length of branches of the tree, we need relays on each branch for assistantship. The number of hops on each branch is $\Theta(\frac{1}{r\sqrt{k}})$. There are totally $k-1$ branches, so the total number of nodes that form the spanning tree is on the order of $\Theta(\frac{\sqrt{k}}{R})$. Since the critical transmission range is $\Theta(\sqrt{\frac{\log(1/R)}{n}})$, according to (1), the per-node capacity is

$$C = \Theta\left(\frac{R}{\sqrt{k}\log(1/R)}\right). \tag{8}$$

If $R = \Omega(\sqrt{\frac{1}{k}})$, namely the transmission range is large enough for nodes on the either sides of branch communicate directly, the overall redundancy is $\Theta(k)$. Therefore,

$$C = \Theta\left(\frac{1}{k\log(1/R)}\right). \tag{9}$$

4.4.2 Delay

The largest number of hops is $\Theta(1/R)$ from the source to each destination. So we only need to determine the average delay on each hop. We treat the procedure that one node transmits its packet to the nodes in adjacent circles as a small i.i.d multicast subnetwork.

If $R = O(\sqrt{\frac{1}{k}})$, the radius of active area is much less than average branch length of original spanning tree so that multihop is needed. In this case, almost all bifurcate points on the original spanning tree cannot communicate directly. Let k' denote the number of receivers for each bifurcate points. Since there are totally $\Theta(k)$ bifurcate points on the original spanning tree, each bifurcate point increases the number of terminals by $k'-1$, so finally there are $\Theta(k\cdot k')$ terminals. Notice that the total number of terminals is $k-1$, we have $k' = \Theta(1)$. This result shows that each bifurcate point has constant number of receivers. We can treat it as unicast problem.

Now that all $k-1$ nodes are deployed uniformly and randomly in the network, it is obvious that there exists at least one destination that is at $\Theta(1)$ far away from the source node. According to Theorem 4.3, it requires $\Theta(1/R)$ hops for the packet to reach that "furthest" destination. Then, we consider delay on each hop d'. According to Theorem 4.3, d' is on the same order as that in a scaled network in [5], by setting cell length as r. Therefore, $d' = \Theta(\frac{nR^2}{\log(1/R)})$. So the total delay is

$$D = \Theta\left(\frac{1}{R}\right)\cdot\Theta\left(\frac{nR^2}{\log(1/R)}\right) = \Theta\left(\frac{nR}{\log(1/R)}\right). \tag{10}$$

If $R = \Omega(\sqrt{\frac{1}{k}})$, the radius of active area is much larger than average length of original spanning tree so that no multihop is needed. Moreover, for each sender, its transmission area can cover a group of bifurcate points in original spanning tree, namely it may cover a subtree of the original spanning tree and the average size of each subtree is $\Theta(kR^2)$. Thus, the sender can send its packet to all other bifurcate points on the subtree, which can form a subnetwork with a multicast flow. The critical transmission range is the same as unicast $r = \Theta(\sqrt{\frac{\log(1/R)}{n}})$. This subnetwork is a scaled one in [10] if we set the cell length as r. The delay on each hop is therefore $\Theta(\frac{n\log(kR^2)}{k\log(1/R)})$. Notice that the number of hops is $\Theta(\frac{1}{R})$, the total delay is

$$D = \Theta\left(\frac{1}{R}\right)\cdot\Theta\left(\frac{n\log(kR^2)}{k\log(1/R)}\right) = \Theta\left(\frac{n\log(kR^2)}{kR\log(1/R)}\right). \tag{11}$$

When $R = \Theta(\sqrt{\frac{\log n}{n}})$, the answer is $\Theta(\sqrt{\frac{n}{\log n}})$. And when $R = \Theta(1)$, the answer is $\Theta(\frac{n\log k}{k})$. The results fit the answers of stationary case [9] and i.i.d model [10]. Specifically, when $k = 1$, the answer also perfectly covers the answer of unicast scheme with LSRM, which covers the answers of stationary unicast [1] and i.i.d unicast [5]. The answer is quite general.

4.5 Data-Copy Scheme

We can use data-copy to reduce delay in mobile networks. The source sends data to a group of relays so that the probability for these relays to reach the destination becomes larger. According to Theorems 3.4 and 3.5, we can get delay and capacity of data-copy scheme by extending the results of non-data-copy scheme given in the previous part.

1. For $R = O(\sqrt{\frac{1}{k}})$, multihop is necessary in linking bifurcate nodes in original spanning tree. And each bifurcate node sends its packet to the constant number of receivers. Therefore, the transmission between each bifurcate node and its receivers can be treated as unicast. According to Theorem 3.4, if the number of data-copy on each hop is m, the per-node capacity is $\Theta(\frac{R}{m\sqrt{k}\log(1/R)})$, and the corresponding delay decreases to

$$\Theta\left(\frac{1}{R}\right)\cdot\Theta\left(\frac{nR^2}{m\log(1/R)}\right) = \Theta\left(\frac{nR}{m\log(1/R)}\right).$$

2. For $R = \Omega(\frac{1}{k})$, each bifurcate node can transmit data to a group of bifurcate node in original spanning tree. Therefore, no relay is needed and we can treat the transmissions between each bifurcate node and its receivers as multicast. According to Theorem 3.5, if the number of total data-copy is m (mR^2 for each subnetwork), the per-node capacity of each submulticast-network is $\Theta(\frac{1}{(mR^2+kR^2)\log(1/R)})$. Since there are totally $\Theta(1/R^2)$ subulticast-networks, the per-node capacity is, therefore, $\Theta(\frac{1}{(m+k)\log(1/R)})$. The corresponding delay is

$$\Theta\left(\frac{1}{R}\right) \cdot \Theta\left(\frac{n\log(kR^2)}{(k+m)\log(1/R)}\right) = \Theta\left(\frac{n\log(kR^2)}{(k+m)R\log(1/R)}\right).$$

The answers can cover that under the stationary model [9] and the i.i.d mobility model [10]. The answer for stationary case is the same as that of non-data-copy scheme, in other words, there is no data-copy scheme in stationary case.

5 Capacity and Delay in GSRM

In this section, we study the performance of GSRM. Under GSRM, all the nodes' displacement in adjacent time slots is less than a certain range. Our results can cover that of i.i.d unicast [5].

We have the following main results in this section:

- For $R = O(\frac{1}{\sqrt{k}})$, non-data-copy scheme, per-node capacity is
$$\Theta\left(\frac{1}{k}\right),$$
and delay is $\Theta(\frac{1}{\sqrt{k}R^2} + \frac{n\log k}{k})$.
- For
$$R = O\left(\frac{1}{\sqrt{m+k}}\right),$$
data-copy scheme, per-node capacity is $\Theta(\frac{1}{m+k})$, and delay is
$$\Theta\left(\frac{1}{\sqrt{m+k}R^2} + \frac{n\log k}{m+k}\right).$$
- For
$$R = \Omega\left(\frac{1}{\sqrt{k}}\right),$$
non-data-copy scheme, per-node capacity is $\Theta(\frac{1}{k})$, and delay is $\Theta(\frac{1}{R} + \frac{n\log k}{k})$.
- For
$$R = \Omega\left(\frac{1}{\sqrt{m+k}}\right),$$
data-copy scheme, per-node capacity is $\Theta(\frac{1}{m+k})$, and delay is $\Theta(\frac{1}{R} + \frac{n\log k}{m+k})$.
- When $k = 1$, the answer is for unicast scheme.

5.1 Critical Transmission Range

The first step is to determine the critical transmission range of network under GSRM. In this part, we show that the critical transmission range is $\Theta(\sqrt{\frac{1}{n}})$.

Theorem 5.1. *For all wireless ad hoc network with n wireless nodes, and inside a unit square network area, the critical transmission range cannot be $o(\sqrt{\frac{1}{n}})$, namely, it is below the order of average distance between adjacent nodes.*

Proof. If there exists a critical transmission range $r = o(\sqrt{\frac{1}{n}})$, we prove that the network is not connected. In fact, for each node, its transmission area is $\pi r^2 = \Theta(r^2)$. Since $r = o(\sqrt{\frac{1}{n}})$, the total area where at least one node can transmit data cannot exceed $\Theta(nr^2) = o(1)$. Therefore, the area where no node can transmit data is on the order of $\Theta(1)$. Thus, nodes are w.h.p situated in this area and cannot make connection with any other node. The network is disconnected. So the critical transmission range cannot below $\Theta(\sqrt{\frac{1}{n}})$.

Then, we prove that the critical transmission range on the order $\Theta(\sqrt{\frac{1}{n}})$ can make the network connected. In fact, when $r = \Theta(\sqrt{\frac{1}{n}})$, the area where a node can transmit data is $\pi r^2 = \Theta(\frac{1}{n})$. Since all the nodes are uniformly distributed in the network area, the probability of the event that other nodes enter this area is $\Theta(\frac{1}{n})$. As a result, the probability of the event that all the other $n-1$ nodes are out of this area is $(1-\Theta(\frac{1}{n}))^{n-1} = \Theta(1) < 1$, which is a constant. When given a certain delay, each node can meet another node inside its transmission area. All nodes can connect with any other node within finite delay, the network is therefore connected. □

So the critical transmission range under GSRM is $\Theta(\sqrt{\frac{1}{n}})$. The main reason is that every node can move to any position under GSRM. The whole network is redeployed in each new time slot. Therefore, the randomness of deployment does not change the connectivity.

5.2 Multicast Scheme, Non-Data-Copy Strategy

A node can reach all the position in the network area. So each packet waits for direct transmission until the sender meet the receiver. Since there can be no extra redundancy except k for each destination node in GSRM, per-node capacity can approach $c = \Theta(\frac{1}{k})$.

Then, we consider delay. In order to simplify the problem, we change our model for a little bit. In Fig. 7, we temporarily assume that the active area for each node is a square with side length $2R$, the center of which is placed on the original point in previous time slot. Therefore, the movement range is restricted by R along both horizontal direction and vertical direction. We assume that the distribution of probability between $[-R, R]$ is uniform along both directions that are mutually independent to each other. Then, we can simply consider the case on horizontal direction can be treated similarly as that on vertical direction.

Lemma 5.1. *Suppose $\psi_1, \psi_2, \ldots, \psi_g$ are random variables, and the probability density functions of them are even and*

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

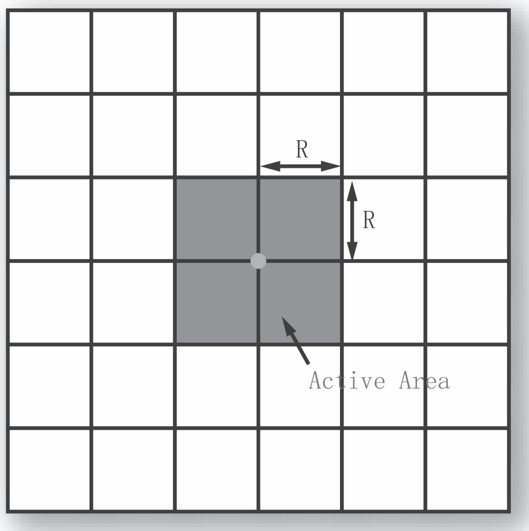


Fig. 7. Squared active area under GSRM.

independent to each other. When g goes to infinity, the variable $\psi = \sum_{i=1}^{g} \psi_i$ *obeys Normal Distribution.*

Proof. Assume the probability density function (PDF) of ψ_i is $f_i(x)$, and its Fourier transform is $F_i(\omega)$. Since $\psi_1, \psi_2, \ldots, \psi_g$ obey to same distribution, $f_i(x)$ and $F_i(\omega)$ keep the same for $i = 1, 2, \ldots, g$. Notice that $F_i(0) = \int_{-\infty}^{+\infty} f_i(x)dx = 1$. We use Taylor's expand, we get $F_i(\omega) = 1 - \theta\omega^2 + o(\omega^2)$, where θ is a positive constant. Assume the PDF of $\psi = \psi_1 + \psi_2 + \cdots + \psi_g$ is $f(x)$ and its Fourier transform is $F(\omega)$. Notice that

$$f(x) = f_1(x) * f_2(x) * \cdots * f_g(x),$$

where operator "$*$" represents convolution operation. Therefore, we have $F(\omega) = F_1(\omega) \cdot F_2(\omega) \cdots F_g(\omega) = (1 - \theta\omega^2 + o(\omega^2))^g \to \exp(-g\theta\omega^2)$. Use Fourier inverse transform, $f(t)$ obeys Normal distribution. □

Theorem 5.2. *Assume initial distance between the sender and receiver is l. Let E denote the event they meet after T time slots. If* $T = o(g)$, *then w.h.p* $P(E) = 0$, *where* $g = \Theta(\frac{l^2}{R^2})$.

Proof. Consider relative movement between X_1 and X_2. First, we can decompose it into two one-dimensional orthogonal movements, and consider one of them. Let ψ_i denote the variation of relative position between two nodes at time slot i. The variance of ψ_i is on the order of $\Theta(R^2)$.

The p.d.f of $\psi = \sum_{i=1}^{g} \psi_i$ is $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$ according to Lemma 5.1, where $\sigma^2 = \Theta(gR^2)$. We require the exponent term of $f(l)$ cannot tend to zero because no polynomial gain can provide such compensation. Then, we need $\frac{l^2}{2\sigma^2} = O(1)$. Therefore, $l = O(\sigma) = O(r\sqrt{g})$. Obviously, $g = \Omega(\frac{l^2}{R^2})$. □

5.2.1 Capacity

Under non-data-copy scheme, all the packets are sent directly to k destinations. So the overall redundancy is k. Note that the critical transmission range for GSRM is $\Theta(\frac{1}{n})$, according to (1), per-node capacity is

$$C = \Theta\left(\frac{1}{k}\right). \tag{12}$$

5.2.2 Delay

We divide the overall delay into two phases: the diffusion phase and the transmission phase, whose delays are D_f and D_t, respectively. In diffusion phase, all the nodes move freely for a period of time in order to make packet holders go further. Within a certain time period $D_f(l)$, each node has an opportunity (in an order sense) to reach any place within range l from its original place. In transmission phase, the packet holders transmit data to new destination nodes.

Theorem 5.3. *For each packet, the diffusion delay to enable this packet reach any place within l* ($l = \Omega(\sqrt{\frac{1}{k}})$) *away from its original place is: If* $R = o(\sqrt{\frac{1}{k}})$, $D_f(l) = \Theta(\frac{l}{\sqrt{k}R^2})$; *If*

$$R = \omega\left(\sqrt{\frac{1}{k}}\right), D_f(l) = \Theta\left(\frac{l}{R}\right).$$

Proof. For $R = o(\sqrt{\frac{1}{k}})$, by Theorem 5.2, it takes at least $\Theta(\frac{1}{kR^2})$ time slots for the source to move to a target $\Theta(\sqrt{\frac{1}{k}})$ far away. Then, the packet can meet destinations. These destinations continue to move until they meet other destinations. Repeat this process until the radius of reachable area for the packet is $\Theta(l)$. Thus, the diffusion time is $\Theta(\frac{l}{\sqrt{k}R^2})$.

For $R = \omega(\sqrt{\frac{1}{k}})$, the radius of reachable area after $\Theta(1)$ time slots is $\Theta(R)$, which contains averagely $\Theta(kR^2)$ destinations. Then, the packet can be transmitted to new destinations $\Theta(R)$ far away. Continue this process, the radius of reachable area becomes $\Theta(l)$ after $\Theta(\frac{l}{R})$ time slots. □

Theorem 5.4. *If m of k destination nodes have received the packet, and all the* unsent destinations[7] *are deployed uniformly, then the average delay for a new receiver to get the packet from packet holders is on the order* $\Theta(\frac{n}{m(k-m)})$.

Proof. Since all the nodes are deployed uniformly in the network area, so that the $k - m$ unsent destinations and the remaining $n - k + m$ nodes are equivalent to each other. Once a packet holder meets another node, the probability that this node is a unsent destination node is $\frac{k-m}{n-1}$. Therefore, the average delay is $\frac{n-1}{k-m} = \Theta(\frac{n}{k-m})$. Notice that there are $m + 1$ packet holders, the average delay is therefore $\Theta(\frac{n}{(m+1)(k-m)})$. Since $m = \Omega(1)$, the theorem is proved. □

Then, consider the order of the overall delay D. First, we consider the lower bound, when we simply assume that the unsent destinations are always deployed uniformly. At the very beginning, none of k destination nodes has received the packet. According to Theorem 5.4, the average delay for the sender to transmit data to the first receiver is $\Theta(\frac{n}{k})$. Then there are two packet holders, the average delay for them to transmit data to a new receiver is $\Theta(\frac{n}{2(k-1)})$. Repeat this process, the expectation of delay should be lower bounded by

7. i.e., destinations which have not obtained the packet.

$$E(D) = \Omega\left(\sum_{i=0}^{k-1} \frac{n}{(1+i)(k-i)}\right) = \Omega\left(\frac{n \log k}{k}\right). \quad (13)$$

The other limitation of delay is that the data packet should have probability to be transmitted to every position in the network after time period of D. According to Theorem 5.3, the overall delay D need to be larger than the diffusion delay $D = \Omega(D_f(1))$. So the lower bound of delay is on the order $\Theta(\max(\frac{n \log k}{k}, D_f(1)))$.

Then, we prove this lower bound is achievable. In fact, we can make the delay of transmission phase as $D_t = \Theta(\frac{n}{k})$. Consider the ith phase, assume m_i destinations of k have received the packet after $i-1$ transmission phases and diffusion phases. Since the average area occupied by each destination is $\Theta(\frac{1}{k})$, the total area that m_i packet holders occupy is $\Theta(\frac{m_i}{k})$. Let l denote the radius of the minimum circular area that covers all these m_i packet holders. The number of unsent destinations that inside this circular area is $\Theta(k(\pi l^2 - \Theta(\frac{m_i}{k})))$. In order to make the unsent destinations deployed uniformly in the network, the density that the unsent destinations are in this circular area should be equal to that in the whole network. So $\pi l^2 / \Theta(k(\pi l^2 - \Theta(\frac{m_i}{k}))) = \Theta(\frac{1}{k})$. By solving this equation, we have $l = \Theta(\sqrt{\frac{m_i}{k}})$. According to Theorem 5.3, the delay cost by the former $i-1$ phases and by the current ith phase should be no less than $D_f(\Theta(\sqrt{\frac{m_i}{k}}))$. According to Theorem 5.4, there are utmost $\Theta(\log k)$ transmission phases required for all k destinations to receive the packet. So the average delay satisfies $E(D) \le D_f(1) + \frac{n \log k}{k}$.

For $R = o(\sqrt{\frac{1}{k}})$, the expectation of delay is

$$\Theta\left(\frac{1}{\sqrt{k}R^2} + \frac{n \log k}{k}\right). \quad (14)$$

For $R = \Omega(\sqrt{\frac{1}{k}})$, the expectation of delay is

$$\Theta\left(\frac{1}{R} + \frac{n \log k}{k}\right). \quad (15)$$

5.3 Multicast Scheme, Data-Copy Strategy

5.3.1 *Capacity*

When there are totally m data copies in the network. The redundancy is $k+m$ instead of k. So the per-node capacity is

$$C = \Theta\left(\frac{1}{k+m}\right). \quad (16)$$

5.3.2 *Delay*

When m-data-copy is used. Similarly, as we did with non-data-copy scheme. Consider the transmission delay and the diffusion delay, respectively. The transmission delay can be calculated by summing up the following series:

$$E(D) = \Omega\left(\sum_{i=0}^{k-1} \frac{n}{(1+i+m)(k-i)}\right) = \Omega\left(\frac{n \log k}{k+m}\right). \quad (17)$$

Then, consider diffusion delay. Since m relay nodes and k destination nodes can help delivering the data packet. According to Theorem 5.3, the diffusion delay for the packet have opportunity to be transmitted to $\Theta(l)$ away from its original place via mutual delivering and mobility is on the order $\Theta(\frac{l}{\sqrt{k+mR^2}})$ for $R = o(\sqrt{\frac{1}{m+k}})$, and $\Theta(\frac{1}{R})$ for

$$R = \Omega\left(\sqrt{\frac{1}{m+k}}\right),$$

respectively. Therefore, we have the following results:

For $R = o(\sqrt{\frac{1}{m+k}})$, the average delay is

$$E(D) = \Theta\left(\frac{1}{\sqrt{m+k}R^2} + \frac{n \log k}{m+k}\right). \quad (18)$$

For $R = \Omega(\sqrt{\frac{1}{m+k}})$, the average delay is

$$E(D) = \Theta\left(\frac{1}{R} + \frac{n \log k}{m+k}\right). \quad (19)$$

6 Discussion

6.1 Delay-Capacity Trade-Off under LSRM

We find that there is a turning point at $R = \Theta(\sqrt{\frac{1}{k}})$ as shown in Fig. 1. The location of this turning point is based on the average length of distance between adjacent destinations. When $R = o(\sqrt{\frac{1}{k}})$, namely mobility speed R is smaller than such length, the whole transmission flow is composed of subunicasts. So the D/C trade-off keeps unchanged; When $R = \omega(\sqrt{\frac{1}{k}})$, namely mobility speed R is large enough to reach a group of destinations, the whole transmission flow is composed of submulticasts, which can reduce D/C trade-off as the scale of submulticasts becomes larger. Optimal D/C trade-off is achieved when $R = \Theta(1)$ as that in i.i.d model.

6.2 Delay-Capacity Trade-Off under GSRM

The curve of D/C trade-off ratio versus R is divided into two cases depending on the relationship between k and n. For $k = o(n^{\frac{2}{3}})$, namely the destinations are distributed sparsely, the turning point is located at $R = \Theta(\frac{k^{0.25}}{\sqrt{n}})$ as shown in Fig. 2. Before this turning point, the diffusion delay for each packet to reach the edge of the network is more significant than the transmission delay. Since the diffusion delay decreases as R increases, thus delay-capacity trade-off is improved; After that point, the transmission delay is more significant. The trade-off keeps unchanged because the transmission delay is independent of the mobility speed R.

For $k = \omega(n^{\frac{2}{3}})$, namely the destinations are distributed densely, the turning point moves to $R = \Theta(\frac{k}{n})$ as shown in Fig. 2. The phenomenon and explanation are familiar with those of the first case, except for another turning point at $R = \Theta(\sqrt{\frac{1}{k}})$, which separates different descending rate on either side. The descending rate on the right side is smaller because the source can reach destinations within constant steps. And the destinations can help spreading the packet to a much larger region than that via the source's movement itself.

6.3 Improvement Region in Mobile Multicast

In multicast network under LSRM, mobility of nodes does not improve the D/C trade-off when the speed is low, until such speed is larger than the turning point shown in Fig. 1. Therefore, mobility is unnecessary when mobility speed is restricted below the turning point.

Under GSRM, the mobility improves the D/C trade-off when the speed is low. When speed reaches the turning point shown in Fig. 2, the D/C trade-off keeps uniform. Thus, we no longer need to increase the mobility speed any more.

We name the region that R improves D/C trade-off as *Improvement Region*. For both LSRM and GSRM, the range of Improvement Regions increases as k increases. Particularly, in unicast, i.e., $k = 1$, the D/C trade-offs always keep stable and never influenced by mobility speed R. In contrast, in broadcast, i.e., $k = n - 1$, the D/C trade-offs are always improved by increasing mobility speed R. Therefore, the impact of mobility speed is more significant when the destinations are distributed more densely in the network.

7 Conclusion

In this paper, we study the relationship between mobility speed R and delay-capacity trade-off ratio. We show that increasing mobility speed R can improve delay-capacity trade-off in multicast network when R is in a specified region. Such region is larger or less than a certain turning point, depending on k and n, which shows that the increasing mobility speed does not always improve delay-capacity trade-off.

We provide a general Speed-Restricted mobility Model, including two particular forms LSRM and GSRM. In different forms, the mobility speed R influences the delay-capacity trade-off in different manners. Both two cases converge to i.i.d when $R = \Theta(1)$, which has optimal D/C trade-off ratio. In both LSRM and GSRM, as k increases, the range of Improvement Region becomes larger so that the impact of mobility is more significant. Such results of MANET can explicitly explain the impact of velocity to distributed system, thus can be utilized to understand the asymptotic achievable capacity when we want to distribute mobile sensors in a region. We have not studied the impact of larger transmission range and the combination of LSRM and GSRM, which could be future works.

Acknowledgments

This work is supported by National Fundamental research grant (2010CB731803, 2009CB3020402), NSF China (No. 60702046, 60832005, 60972050, 60632040); China Ministry of Education (No. 20070248095); China Ministry of Education Fok Ying Tung Fund (No. 122002); Qualcomm Research Grant; China International Science and Technology Cooperation Programm (No. 2008DFA11630); PUJIANG Talents (08PJ14067); Shanghai Innovation Key Project (08511500400); National Key Project of China (2009ZX03003-006-03, 2009ZX03002-003, 2009ZX03002-005); and National High tech grant of China (2009AA01Z248, 2009AA011802).

References

[1] P. Gupta and P.R. Kumar, "The Capacity of Wireless Networks," *IEEE Trans. Information Theory,* vol. 46, no. 2, pp. 388-404, Mar. 2000.

[2] M. Grossglauser and D.N.C. Tse, "Mobility Increases the Capacity of Ad Hoc Wireless Networks," *IEEE/ACM Trans. Networking,* vol. 10, no. 4, pp. 477-486, Aug. 2002.

[3] N. Bansal and Z. Liu, "Capacity, Delay and Mobility in Wireless Ad-Hoc Networks," *Proc. IEEE INFOCOM,* Apr. 2003.

[4] E. Perevalov and R. Blum, "Delay Limited Capacity of Ad Hoc Networks: Asymptotically Optimal Transmission and Relaying Strategy," *Proc. IEEE INFOCOM,* Apr. 2003.

[5] M.J. Neely and E. Modiano, "Capacity and Delay Tradeoffs for Ad Hoc Mobile Networks," *IEEE Trans. Information Theory,* vol. 51, no. 6, pp. 1917-1937, June 2005.

[6] S. Toumpis and A.J. Goldsmith, "Large Wireless Networks under Fading, Mobility, and Delay Constraints," *Proc. IEEE INFOCOM,* Mar. 2004.

[7] X. Lin and N.B. Shroff, "The Fundamental Capacity-Delay Tradeoff in Large Mobile Wireless Networks," technical report, http://cobweb.ecn.purdue.edu/linx/papers.html, 2004.

[8] G. Sharma and R. Mazumdar, "Scaling Laws for Capacity and Delay in Wireless Networks with Random Mobility," *Proc. IEEE Int'l Conf. Comm. (ICC),* 2004.

[9] X. Li, S. Tang, and O. Frieder, "Multicast Capacity for Large Scale Wireless Ad Hoc Networks," *Proc. ACM MobiCom,* Sept. 2007.

[10] C. Hu, X. Wang, and F. Wu, "MotionCast: On the Capacity and Delay Tradeoffs," *Proc. ACM MobiHoc,* May 2009.

[11] A. Keshavarz-Haddad, V. Ribeiro, and R. Riedi, "Broadcast Capacity in Multihop Wireless Networks," *Proc. ACM MobiCom,* Sept. 2006.

[12] P. Jacquet and G. Rodolakis, "Multicast Scaling Properties in Massively Dense Ad Hoc Networks," *Proc. Int'l Conf. Parallel and Distributed Systems,* July 2005.

[13] S. Shakkottai, X. Liu, and R. Srikant, "The Multicast Capacity of Large Multihop Wireless Networks," *Proc. ACM MobiHoc,* Sept. 2007.

[14] A.E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-Delay Trade-Off in Wireless Networks," *Proc. IEEE INFOCOM,* Mar. 2004.

[15] M. Zhao and W. Wang, "A Novel Semi-Markov Smooth Mobility Model for Mobile Ad Hoc Networks," *Proc. IEEE Global Telecomm. Conf. (GLOBECOM),* Nov. 2006.

[16] G. Sharma, R. Mazumdar, and N.B. Shroff, "Delay and Capacity Trade-Offs in Mobile Ad Hoc Networks: A Global Perspective," *IEEE Trans. Networking,* vol. 15, no. 5, pp. 981-992, Oct. 2007.

[17] X. Lin, G. Sharma, R. Mazumdar, and N.B. Shroff, "Degenerate Delay-Capacity Trade-Offs in Ad Hoc Networks with Brownian Mobility," *Joint Issue of IEEE Trans. Networking and Information Theory,* vol. 52, no. 6, pp. 2777-2784, June 2006.

[18] G. Sharma and R. Mazumdar, "On Achievable Delay/Capacity Trade-Offs in Mobile Ad Hoc Networks," *Proc. Workshop Modeling and Optimization in Ad Hoc Mobile Networks,* 2004.

[19] L. Ying, S. Yang, and R. Srikant, "Optimal Delay-Throughput Trade-Offs in Mobile Ad Hoc Networks," *IEEE Trans. Information Theory,* vol. 54, no. 9, pp. 4119-4143, Sept. 2008.

[20] D. Ciullo, V. Martina, M. Garetto, and E. Leonardi, "Impact of Correlated Mobility on Delay-Throughput Performance in Mobile Ad-Hoc Networks," *Proc. IEEE INFOCOM,* Mar. 2010.

[21] M. Garetto, P. Giaccone, and E. Leonardi, "Capacity Scaling in Ad Hoc Networks with Heterogeneous Mobile Nodes: The Sub-Critical Regime," *ACM/IEEE Trans. Networking,* vol. 17, no. 6, pp. 1888-1901, Dec. 2009.

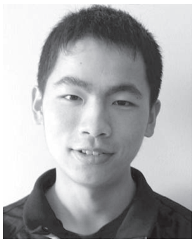
[22] M. Steele, "Growth Rates of Euclidean Minimal Spanning Trees with Power Weighted Edges," *The Annals of Probability,* vol. 2, no. 16, pp. 1767-1787, 1988.

[23] X. Li, Y. Liu, S. Li, and S. Tang, "Multicast Capacity of Wireless Ad Hoc Networks under Gaussian Channel Model," *IEEE/ACM Trans. Networking,* vol. 18, no. 4, pp. 1145-1157, Aug. 2010.

[24] S. Li, Y. Liu, and X. Li, "Capacity of Large Scale Wireless Networks under Gaussian Channel Model," *Proc. ACM MobiCom,* Sept. 2008.

Xinbing Wang received the BS degree (with hons) in automation from Shanghai Jiaotong University, China, in 1998, the MS degree in computer science and technology from Tsinghua University, Beijing, China, in 2001, and the PhD degree, major in the electrical and computer engineering, minor in the mathematics from North Carolina State University, Raleigh, in 2006. Currently, he is an associate professor in the Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai, China. His research interests include resource allocation and management in mobile and wireless networks, TCP asymptotics analysis, wireless capacity, cross-layer call admission control, asymptotics analysis of hybrid systems, and congestion control over wireless ad hoc and sensor networks. He has been a member of the organizing committees of several conferences including IEEE INFOCOM 2011, IEEE ICC 2011, and IEEE Globecom 2011. He is a member of the IEEE.

Yuanzhe Bei received the BE degree in electronic engineering from Shanghai Jiao Tong University, China, in 2009. His research interests are in the area of asymptotic analysis of capacity and connectivity in wireless ad hoc network. He is a student member of the IEEE.



Qiuyu Peng is currently working toward the BE degree in electronic engineering in Shanghai Jiao Tong University, China, and working with Prof. Xinbing Wang in Institute of Wireless Communication Technology of Shanghai Jiao Tong University. His research interests are in the area of asymptotic analysis of capacity and connectivity in wireless ad hoc network. He is a student member of the IEEE.

Luoyi Fu received the BE degree in electronic engineering from Shanghai Jiao Tong University, China, in 2009. She is currently working with Prof. Xinbing Wang toward the master's degree in the Department of Electronic Engineering in Shanghai Jiao Tong University. Her research of interests are in the area of scaling laws analysis in wireless networks. She is a student member of the IEEE.