

# Degenerate Delay-Capacity Trade-offs in Ad Hoc Networks with Brownian Mobility

Xiaojun Lin, Gaurav Sharma, Ravi R. Mazumdar, and Ness B. Shroff

**Abstract**— There has been significant recent interest within the networking research community to characterize the impact of mobility on the capacity and delay in mobile ad hoc networks. In this paper, the fundamental trade-off between the capacity and delay for a mobile ad hoc network under the Brownian motion model is studied. It is shown that the 2-hop relaying scheme proposed by Grossglauser and Tse (2001), while capable of achieving a per-node throughput of  $\Theta(1)$ , incurs an expected packet delay of  $\Omega(\log n/\sigma_n^2)$ , where  $\sigma_n^2$  is the variance parameter of the Brownian motion model. It is then shown that an attempt to reduce the delay beyond this value results in the throughput dropping to its value under static settings. In particular, it is shown that under a large class of scheduling and relaying schemes, if the mean packet delay is  $O(n^\alpha/\sigma_n^2)$ , for any  $\alpha < 0$ , then the per-node throughput must be  $O(1/\sqrt{n})$ . This result is in sharp contrast to other results that have recently been reported in the literature.

## I. INTRODUCTION

Since the seminal work of Gupta and Kumar [1], there has been a lot of interest in characterizing the capacity region of ad hoc networks. A major contribution in this direction was made in [2], where the authors showed that mobility can significantly increase the traffic carrying capacity of an ad hoc network. In particular, the authors proposed a 2-hop relaying scheme, and showed that it can achieve a per-node throughput of  $\Theta(1)$ <sup>1</sup>. However, delay related issues were not addressed in [2]. In fact, it was pointed out in [2] that their 2-hop relaying scheme could potentially incur an unbounded delay.

There has been substantial recent work on the joint characterization of delay and capacity in mobile ad hoc networks [3]–[8]. The type of node mobility studied in the literature includes the so-called *i.i.d* mobility [4], [7], [8], random way-point mobility [5], [6], Brownian motion [3], [6], and Markovian mobility [4]. The results in these works are of a similar flavor.

Xiaojun Lin, Gaurav Sharma, and Ness B. Shroff are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907, USA (email: {linx.gsharma,shroff}@ecn.purdue.edu).

Ravi R. Mazumdar is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (email: mazum@ece.uwaterloo.ca).

<sup>1</sup>We use the following notation throughout:

$$\begin{aligned} f(n) = o(g(n)) &\leftrightarrow \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0, \\ f(n) = O(g(n)) &\leftrightarrow \limsup_{n \rightarrow \infty} \frac{f(n)}{g(n)} < \infty, \\ f(n) = \omega(g(n)) &\leftrightarrow g(n) = o(f(n)), \\ f(n) = \Omega(g(n)) &\leftrightarrow g(n) = O(f(n)), \\ f(n) = \Theta(g(n)) &\leftrightarrow f(n) = O(g(n)) \text{ and } g(n) = O(f(n)). \end{aligned}$$

It is first shown that one can achieve a per-node throughput of  $\Theta(1)$  with a bounded average delay, using scheduling schemes that are variants of Grossglauser-Tse 2-hop relaying scheme. These works then report *trade-offs* between the capacity and delay; i.e., the delay can be reduced if one is willing to accept a lower per-node throughput, and vice versa. The capacity-delay trade-offs are achieved either by means of introducing some redundancy in the 2-hop relaying scheme [4]–[6], or by adjusting the cell size [3], or both [7], [8]. All previous works in the literature have reported a “smooth” capacity-delay trade-off under their respective settings. For instance, under the random way-point mobility model, the authors of [5] report a scheme that can achieve  $\Theta(n^{\alpha-1})$  per-node throughput at  $\Theta(n^\alpha)$  delay, for any  $\alpha \in [1/2, 1]$ . Hence, by reducing  $\alpha$  to  $\alpha - \epsilon$ ,  $\epsilon > 0$ , one can reduce the delay by a factor of  $\Theta(n^\epsilon)$ , at the cost of reducing the capacity by the same factor. Similar type of smooth trade-offs have been reported under the *i.i.d.* mobility model as well [4], [7], [8].

In this paper, we study the trade-off between the capacity and delay under the Brownian motion model [3], [6]. We show that there is *virtually no trade-off between the capacity and delay under the Brownian motion model* (see Fig. 1). In particular, we show that under a large class of scheduling and relaying schemes, in order to achieve a delay of  $\Theta(n^\alpha/\sigma_n^2)$  for any  $\alpha < 0$ , where  $\sigma_n^2$  is the variance parameter of Brownian motion, the per-node throughput must be  $O(1/\sqrt{n})$ . Further, we show that the 2-hop relaying scheme proposed by Grossglauser and Tse [2], while capable of achieving a per-node throughput of  $\Theta(1)$ , incurs an expected packet delay of  $\Omega(\log n/\sigma_n^2)$ . Note that a per-node throughput of  $\Theta(1/\sqrt{n})$  can be achieved in case of static wireless networks, using multi-hop scheduling scheme [9], [10]. Thus, in order to achieve any capacity gains by exploiting node motion, one must be ready to tolerate huge delays, roughly on the order of  $\Theta(1/\sigma_n^2)$ , which is close to the delay at a per-node throughput of  $\Theta(1)$ . Interestingly, earlier studies of the delay-capacity trade-off under the Brownian motion model were incorrect in that they reported the existence of a smooth trade-off [3], [6].

We note that the results of this paper also apply to other related mobility models such as the Markovian mobility model of [4]. This is because the Brownian motion model can be viewed as a limiting case of these other mobility models. Thus, one would expect the delay-capacity trade-off to be degenerate under these models as well.

We summarize the main contributions of this paper below:

- We rigorously show that for a large class of scheduling and relaying schemes, the achievable capacity-delay trade-off under the Brownian motion model is *degenerate*.

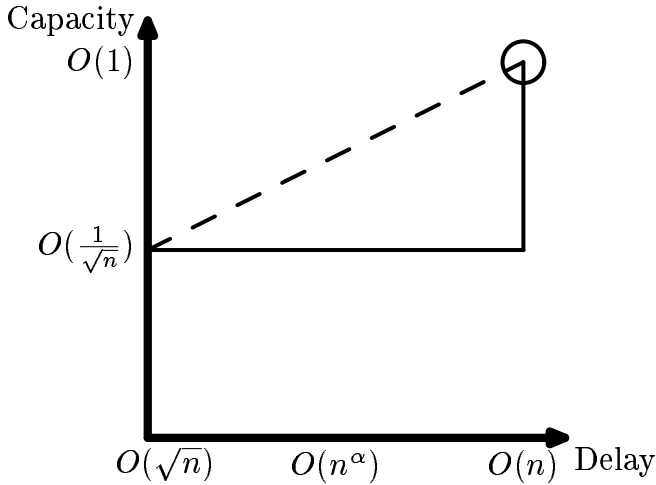


Fig. 1. The degenerate delay-capacity trade-off under the Brownian motion model (the solid line) compared with the “smooth” delay-capacity trade-off under the random way-point mobility model (the dashed line) reported in [5]. We have chosen  $\sigma_n^2 = 1/n$  and ignored all logarithmic terms in the figure.

More precisely, an attempt to achieve a higher per-node throughput than that of static ad hoc networks, results in delays becoming excessively large, roughly on the order of  $\Theta(1/\sigma_n^2)$ .

- We consider the class of generalized 2-hop relaying schemes and show that they incur a delay of  $\Omega(\log(1/a_n)/\sigma_n^2)$ , where  $a_n$  is related to the concept of capture neighborhood and forwarding neighborhood (see Section V). Most scheduling schemes studied in the literature fall into this class with  $a_n = O(n^\alpha)$ , for some  $\alpha < 0$ . Hence, the delay they incur is no less than  $\Omega(\log n/\sigma_n^2)$ . A special case is the 2-hop relaying scheme of [2] which achieves a per-node throughput of  $\Theta(1)$ , and incurs an expected packet delay of  $\Omega(\log n/\sigma_n^2)$ .

It is interesting to compare our results for the Brownian motion model with the results in [5] for the random way-point mobility model. Note that both these models are continuous mobility models (i.e., the motion of the nodes is continuous), and both preserve the uniform distribution of nodes at all times, i.e., an initial uniform distribution of nodes implies that the nodes remain uniformly distributed at all times. However, the capacity-delay relationship under these two models is significantly different. In particular, there exists a smooth trade-off between the delay and capacity under the random way-point mobility model, whereas there is virtually no trade-off under the Brownian motion model. We believe that this difference is a revelation of the fundamental difference in the mobility pattern under these two models. In the random way-point mobility model, nodes move “purposefully,” i.e., during each trip, a node has some target position in mind (chosen uniformly on the sphere) and it moves along a straight-line path, with no “wandering” at all. Thus, the nodes can cover large distances in relatively short time under the random way-point mobility model. This is in contrast to the Brownian motion model, where the nodes always wander around like “drunkards,” staying in a local neighborhood for large duration

of time. It is therefore intuitive to believe that reducing the mobility delay under the Brownian motion model would be more difficult.

The rest of the paper is organized as follows. In the next section, we describe our network model and the Brownian motion model, followed by some basic properties of the Brownian motion model in Section III. We then derive our main result in Section IV, showing that there is virtually no trade-off under the Brownian motion model. In Section V, we analyze the delay performance of generalized 2-hop relaying schemes. We end this paper with some concluding remarks in Section VII.

## II. SYSTEM MODEL

We consider an ad hoc network, consisting of  $n$  mobile nodes. For ease of exposition, we consider two different network shapes: a unit sphere and a unit square in  $\mathbb{R}^2$ . The choice between the two will be made based on technical convenience. Note that the scaling laws for capacity and delay are the same for both the above network shapes. In fact, they are the same for any connected, closed, and convex network shape of unit area in  $\mathbb{R}^2$ . The initial distribution of the nodes is assumed to be uniform. Under the Brownian motion model that we consider, an initial uniform distribution of nodes implies that the nodes remain uniformly distributed at all times.

For simplicity, we assume that each node, say node  $i$ , communicates with a single destination node, say node  $d(i)$ , and that the mapping  $i \mapsto d(i)$  is bijective. We assume a uniform traffic pattern, i.e., each *source* generates traffic at the same rate of  $\lambda$  bits per second for its destination node. We further assume that the packet arrival processes at each node is independent of the node mobility process. The communication between any *source-destination* pair can possibly be via multiple other nodes acting as relays. That is, the source node could, if possible, send a packet directly to the destination node; or it could forward the packet to one or more relay nodes; the relay nodes could themselves forward the packet to other relay nodes; and finally a relay node or the source node itself could deliver the message to the destination node.

To model the effect of interference, we use the Protocol Model of [1]. Let  $W$  be the bandwidth of the system in bits per second. Let  $X_t^i$  denote the position of the node  $i$ , for  $i = 1 \dots n$ , at time  $t$ . Node  $i$  can communicate directly with another node  $j$  at the rate of  $W$  bits per second at time  $t$ , if and only if the following interference constraint is satisfied: [1]:

$$d(X_t^k, X_t^j) \geq (1 + \Delta)d(X_t^i, X_t^j) \quad (1)$$

for every other node  $k \neq i, j$  that is simultaneously transmitting. Here,  $\Delta$  is some positive number, and  $d(x, y)$  denotes the Euclidean distance between points  $x, y \in \mathbb{R}^3$ . Note that when the unit of information transmitted is a packet, the above interference constraint must be satisfied over the entire duration of the packet transmission from node  $i$  to node  $j$ .

Let  $S$  denote the surface of the unit sphere, centered at the origin. We assume that nodes move independently on  $S$  according to a Brownian motion model as in [6]. (A similar

Brownian motion model on a 2-d torus is also considered in [3].) It is easier to describe the motion of each node using the spherical coordinates. Let  $\theta_t$  and  $\phi_t$  denote the colatitude and longitude, respectively, of the position of a particular node at time  $t$  ( $0 \leq \theta_t \leq \pi$  and  $0 \leq \phi_t < 2\pi$ ). When a node moves according to the Brownian motion model on the unit sphere  $S$ , the (Itô) stochastic differential equations for the process  $(\theta_t, \phi_t)$  are given by [11]:

$$d\theta_t = \sigma_n dB_t + \frac{\sigma_n^2}{2 \tan \theta_t} dt, \quad (2)$$

and

$$d\phi_t = \frac{\sigma_n}{\sin \theta_t} dB'_t, \quad (3)$$

where  $B_t$  and  $B'_t$  are independent standard one-dimensional Brownian motions (i.e., with variance 1). We call  $\sigma_n^2$  the *variance* of Brownian motion described in (2) and (3). We note that  $\sigma_n^2$  is related to the average time required by a node to move to different parts of the network. A large value of  $\sigma_n^2$  implies that the node will take a short amount of time to move to different parts of the network, and vice versa (see Lemma 1). For analysis, it is useful to project each node's position on the  $z$ -axis. Substituting  $Y_t = \cos \theta_t$  in (2), and using Itô's Lemma, we obtain

$$dY_t = -\sigma_n^2 Y_t dt - \sigma_n \sqrt{1 - Y_t^2} dB_t. \quad (4)$$

Note that  $Y_t$  is a diffusion process with *drift coefficient*  $-\sigma_n^2 Y_t$  and *diffusion coefficient*  $\sigma_n^2(1 - Y_t^2)$ .

### III. BASIC PROPERTIES OF SPHERICAL BROWNIAN MOTION

In this section, we summarize some basic properties of spherical Brownian motion, to be used later in the analysis. We note that all of the following results can also be derived for a Brownian motion on a 2-d torus (square with a wrap around), but the derivations are much more tedious.

Let us consider the motion of a single node. Let  $X_t$  denote its position at time  $t$ , which can be represented using the spherical coordinates  $(\theta_t, \phi_t)$ . Let  $Y_t = \cos \theta_t$  be the projection of the node's position on the  $z$ -axis, and recall that  $Y_t$  is governed by (4).

In what follows, we will need the following result from [11] concerning the expected travel time of  $Y_t$ :

*Lemma 1:* Let  $-1 < a < x < 1$ . Then, in traveling from  $x$  to  $a$ ,  $Y_t$  takes an expected time  $V_a(x)$  given by:

$$V_a(x) = \frac{2}{\sigma_n^2} \log \left( \frac{1+x}{1+a} \right).$$

#### A. The First Hitting Time

The first concept we study is the *first hitting time*. Let  $A$  be an arbitrary region on the sphere. We have the following definition:

*Definition 1:* The first hitting time of  $A$ , denoted by  $T_A$ , is the first time instant at which  $X_t$  enters  $A$ ; i.e.,  $T_A = \inf\{t \geq 0 : X_t \in A\}$ .

Let  $\Pi$  denote the uniform distribution on the unit sphere  $S$ , and denote by  $\mathbb{E}_\Pi$  the expectation conditioned on  $X_0$  being

distributed according to  $\Pi$ . Let  $A = \{x \in S : d_S(x, y) \leq a_n\}$ , where  $y$  is an arbitrary point on  $S$ ,  $d_S$  denotes the shortest geodesic distance on the sphere, and  $a_n > 0$ . For  $a_n \downarrow 0$ , as  $n \rightarrow \infty$ , we have the following result:

*Lemma 2:*  $\mathbb{E}_\Pi[T_A] = \Theta(\log(1/a_n)/\sigma_n^2)$ .

*Proof:* In view of the symmetry of  $S$ , taking  $y$  to be the south pole (i.e., the bottom most point of  $S$  that corresponds to  $\theta = \pi$ ) entails no loss of generality. Now, for  $x \in A$ , we have  $\mathbb{E}[T_A | X_0 = x] = 0$ . For  $x \notin A$ , let  $z_x$  denote its  $z$ -coordinate. Note that the radius of  $S$  is  $\frac{1}{2\sqrt{\pi}}$ . The first time that  $X_t$  enters  $A$  is also the first time that  $Y_t$  travels from  $z_x$  to  $-\cos(2a_n\sqrt{\pi})$ . Using Lemma 1, we obtain

$$\mathbb{E}[T_A | X_0 = x] = \frac{2}{\sigma_n^2} \log \left( \frac{1+z_x}{1-\cos(2a_n\sqrt{\pi})} \right).$$

Integrating over all possible positions of the point  $x$  on  $S$ , and using the fact that  $x$  is uniformly distributed on  $S$ , we obtain

$$\mathbb{E}_\Pi[T_A] = \int_{\theta=2\sqrt{\pi}a_n}^{\pi} \frac{\sin \theta}{\sigma_n^2} \log \left( \frac{1+\frac{\cos \theta}{2\sqrt{\pi}}}{1-\cos(2a_n\sqrt{\pi})} \right) d\theta,$$

and the result follows after straightforward calculations. ■

*Remark 1:* If  $a_n = n^\alpha$ ,  $\alpha < 0$ , then by Lemma 2, the first hitting time is always  $\Theta(\log n/\sigma_n^2)$ , regardless of the value of  $\alpha$ . Even if we take  $a_n = \sqrt{\pi}/4$ , which means that the set  $A$  covers about half of  $S$ , the first hitting time is still  $\Theta(1/\sigma_n^2)$ . Hence, the first hitting time changes very little when the size of the set  $A$  is increased. This result reveals the fundamental difference between the mobility pattern under the Brownian motion model and that under other mobility models (such as the *i.i.d* mobility model [8] and the random way-point mobility model [5]). In these other models, the first hitting time for a set  $A$  decreases substantially when the size of the set  $A$  is increased. On the other hand, Lemma 2 is not completely surprising given the fact that, under the Brownian motion model, the node always wanders around like a “drunkard.” Therefore, it is very difficult for the node to move towards any given destination.

#### B. The First Exit Time

The second concept that we study is the *first exit time*.

*Definition 2:* Let  $A = \{x \in S : d_S(x, y) \leq a_n\}$ . The first exit time for the region  $A$ , denoted by  $\tau_A$ , is the first instant of time at which the Brownian motion started at  $y$  (the center of  $A$ ) exits  $A$ , i.e.,

$$\tau_A = \inf\{t \geq 0 : X_0 = y, X_t \notin A\}.$$

Assuming  $a_n \rightarrow 0$  as  $n \rightarrow \infty$ , we have the following result:

*Lemma 3:*  $\mathbb{E}[\tau_A] = \Theta(a_n^2/\sigma_n^2)$ .

*Proof:* In view of the symmetry of  $S$ , we can set  $y$  to be the north pole of  $S$  (i.e., the top most point of  $S$  that corresponds to  $\theta = 0$ ). It then follows that  $\mathbb{E}[\tau_A]$  is the expected travel time of  $Y_t$  from 1 to  $\cos(2a_n\sqrt{\pi})$ . Using Lemma 1 and performing some straightforward calculations, the result follows. ■

*Remark 2:* From the above discussion it is clear that under the Brownian motion model a node requires  $\Theta(a_n^2/\sigma_n^2)$  time to move a radial distance of  $a_n$ . Thus, the time a Brownian

motion process spends in a region is proportional to the area of the region. This also points to the well known result that the Brownian motion paths are nowhere differentiable [12, p380]. Hence, it is inappropriate to define the “velocity” of a node that is moving in accordance with the Brownian motion model.

#### IV. THE DEGENERATE CAPACITY-DELAY TRADE-OFF

In this section, we show that there is virtually no trade-off between the delay and capacity under the Brownian motion model. Specifically, we will show that whenever the delay constraint is  $O(n^\alpha/\sigma_n^2)$  for any  $\alpha < 0$ , the per-node capacity is  $O(1/\sqrt{n})$ . For ease of exposition, we will be using a planar Brownian motion model in this section. Nonetheless, as we will argue later, the results also hold under the spherical Brownian motion model.

Consider  $n$  nodes on a unit square centered at the origin, executing independent two-dimensional Brownian motions within the square. As will become clear shortly, our results do not depend on how the boundary condition is handled: the Brownian motion could either be reflected at the boundary, or wrap around the boundary (like the 2-d torus model in [3]).

In order to prove the main result of this section, namely the delay-capacity trade-off under the Brownian motion model is degenerate, we need some supporting results (Lemma 4 and Lemma 5 below). The main idea is as follows: If the delay is  $O(n^\alpha/\sigma_n^2)$  for  $\alpha < 0$ , then we can show that the contribution of node mobility in the packet delivery is likely very small. Hence, in order for the packet to be delivered to its destination node, relaying over order  $\Theta(1)$  distance is required; in which case, the achievable per-node throughput can be shown to be  $O(1/\sqrt{n})$ .

We start by showing that, if the delay is  $O(n^\alpha/\sigma_n^2)$  for  $\alpha < 0$ , then the contribution of node mobility in the packet delivery is likely very small. Let  $\mathbf{SQ}(c_n)$  be the square centered at the origin with length  $c_n$  (see Fig. 2). Suppose there are  $k_n \leq n$  nodes, starting at the origin at time 0. Each node then moves according to a two-dimensional Brownian motion with variance  $\sigma_n^2$ , which can be viewed as the composition of two independent one-dimensional Brownian motion along the x-axis and the y-axis, respectively, each having a variance of  $\sigma_n^2/2$ . Let  $p_{k_n}(c_n, t_n)$  denote the probability of the event that one or more of the  $k_n$  nodes ever exit the square  $\mathbf{SQ}(c_n)$  within time  $t_n$ . We have the following result concerning  $p_{k_n}(c_n, t_n)$ :

*Lemma 4:* If there exists  $N_0 < \infty$  such that

$$\frac{c_n^2}{t_n} \geq 8\sigma_n^2 \log n, \text{ for } n \geq N_0, \quad (5)$$

then

$$\lim_{n \rightarrow \infty} p_{k_n}(c_n, t_n) = 0.$$

The following Corollary is an immediate consequence of Lemma 4.

*Corollary 1:* If

$$\liminf_{n \rightarrow \infty} c_n \log n = c > 0$$

$$\limsup_{n \rightarrow \infty} \frac{\sigma_n^2 t_n}{n^\alpha} = c' < +\infty, \text{ for some } \alpha < 0,$$

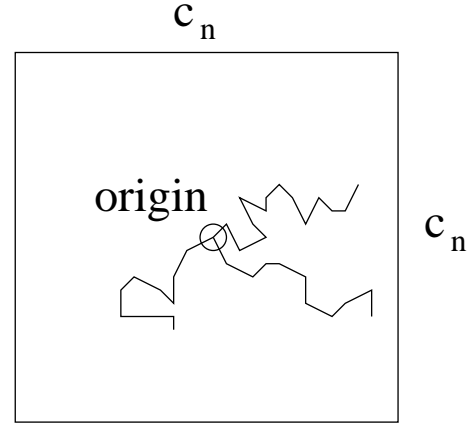


Fig. 2.  $k_n$  nodes starting the origin and executing independent Brownian walks.

then

$$\lim_{n \rightarrow \infty} p_{k_n}(c_n, t_n) = 0.$$

*Remark 3:* Corollary 1 shows that, within  $O(n^\alpha/\sigma_n^2)$  time ( $\alpha < 0$ ), none of the  $k_n$  nodes can possibly travel a  $\Theta(1/\log n)$  distance in *any* direction.

*Proof:* [Proof of Lemma 4] Consider an arbitrary node in the network. Let  $X_t$  be its position at time  $t$ . Let  $B_t^x$  and  $B_t^y$  denote its x-coordinate and y-coordinate, respectively. Then,  $B_t^x$  and  $B_t^y$  are independent one-dimensional Brownian motions with variance  $\sigma_n^2/2$ . Let  $p(c_n, t_n)$  be the probability that this particular node ever exits the square  $\mathbf{SQ}(c_n)$  within time  $t_n$ . Let

$$\begin{aligned} \tau_x^+ &\triangleq \inf\{t \geq 0 : B_t^x = c_n/2\}, \\ \tau_x^- &\triangleq \inf\{t \geq 0 : B_t^x = -c_n/2\}, \end{aligned}$$

and let  $\tau_y^+, \tau_y^-$  be similarly defined with  $B_t^y$  in place of  $B_t^x$ . Using the union bound, and appealing to the symmetry of two-dimensional Brownian motion, we obtain

$$\begin{aligned} p(c_n, t_n) &\leq \mathbf{P}\{\tau_x^+ \leq t_n \text{ or } \tau_x^- \leq t_n \text{ or } \tau_y^+ \leq t_n \text{ or } \tau_y^- \leq t_n\} \\ &\leq 4\mathbf{P}\{\tau_x^+ \leq t_n\}. \end{aligned}$$

Further, using the Reflection Principle for one-dimensional Brownian motion [12, p394], we have

$$\mathbf{P}\{\tau_x^+ \leq t_n\} = 2\mathbf{P}\{B_{t_n}^x \geq c_n/2\}.$$

Since the distribution of  $B_{t_n}^x$  is Gaussian with zero mean and variance  $\sigma_n^2 t_n/2$ , we have,

$$\mathbf{P}\{\tau_x^+ \leq t_n\} = 2 \int_{\frac{c_n}{\sqrt{2\sigma_n^2 t_n}}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du.$$

Using the inequality,

$$\begin{aligned} \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du &\leq \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \frac{u}{x} \exp\left(-\frac{u^2}{2}\right) du \\ &= \frac{1}{\sqrt{2\pi}x} \exp\left(-\frac{x^2}{2}\right), \end{aligned}$$

we have,

$$\mathbf{P}\{\tau_x^+ \leq t_n\} \leq 2 \sqrt{\frac{\sigma_n^2 t_n}{\pi c_n^2}} \exp\left[-\frac{c_n^2}{4\sigma_n^2 t_n}\right].$$

Using (5), we have

$$\mathbf{P}\{\tau_x^+ \leq t_n\} \leq \frac{1}{\sqrt{2\pi \log n}} \exp(-2 \log n) = \frac{1}{n^2 \sqrt{2\pi \log n}}.$$

Hence,

$$p(c_n, t_n) \leq \frac{4}{n^2 \sqrt{2\pi \log n}}.$$

Finally, since there are  $k_n$  nodes, each of them moves according to a two-dimensional Brownian Motion, we have

$$p_{k_n}(c_n, t_n) \leq k_n p(c_n, t_n) \leq \frac{4k_n}{n^2 \sqrt{2\pi \log n}}.$$

Noting that  $k_n \leq n$ , the result follows.  $\blacksquare$

By Corollary 1, if the delay is  $O(n^\alpha / \sigma_n^2)$  for  $\alpha < 0$ , then the contribution of node mobility in the packet delivery is likely very small ( $O(1/\log n)$ ). Hence, in order for the packet to be delivered to its destination node, relaying over order  $\Theta(1)$  distance is required. We now show that if packets are, on an average, relayed over  $\Theta(1)$  distance, then the per-node throughput must be  $O(1/\sqrt{n})$ .

Consider a large enough time interval  $\mathcal{T}$ . The total number of packets communicated end-to-end between all source-destination pairs during the interval is then  $c_p \lambda n \mathcal{T}$ , where  $1/c_p$  is the number of bits per packet. Let  $h_p$  be the number of times the packet  $p$  is relayed, and let  $l_p^h$ , for  $h = 1, \dots, h_p$ , denote the transmission range for the  $h$ -th relaying. We have the following result:

*Lemma 5:* Suppose that there exists a constant  $c > 0$  such that, on an average, packets are relayed over a total distance no less than  $c$ , i.e.,

$$\frac{\sum_{p=1}^{c_p \lambda n \mathcal{T}} \sum_{h=1}^{h_p} l_p^h}{c_p \lambda n \mathcal{T}} \geq c, \quad (6)$$

then

$$\lambda \leq O(1/\sqrt{n}).$$

*Proof:* We use  $d(x, y)$  to denote the Euclidean distance between positions  $x$  and  $y$  within the unit square. Let  $X^i$  denote the position of node  $i$ , for  $i = 1, \dots, n$ . Consider nodes  $i, j$  transmitting directly to nodes  $k$  and  $l$ , respectively, at time  $t$ . Then, under the Protocol Model, in order for the transmissions to be successful, the following inequalities must hold at the time of transmission:

$$\begin{aligned} d(X^j, X^k) &\geq (1 + \Delta) d(X^i, X^k) \\ d(X^i, X^l) &\geq (1 + \Delta) d(X^j, X^l). \end{aligned}$$

Hence,

$$\begin{aligned} d(X^j, X^i) &\geq d(X^j, X^k) - d(X^i, X^k) \\ &\geq \Delta d(X^i, X^k). \end{aligned}$$

Similarly,

$$d(X^i, X^j) \geq \Delta d(X^j, X^l).$$

Therefore,

$$d(X^i, X^j) \geq \frac{\Delta}{2} (d(X^i, X^k) + d(X^j, X^l)).$$

That is, disks of radius  $\frac{\Delta}{2}$  times the transmission range centered at the transmitter are disjoint from each other<sup>2</sup>. We can therefore measure the radio resources that each transmission consumes by the areas of these disjoint disks. Note that the total area of the square is 1; for each of these disks, at least 1/4 of it must lie inside the unit square; and each relaying of a packet lasts  $\frac{1}{c_p W}$  amount of time. Thus,

$$\frac{1}{4} \sum_{p=1}^{c_p \lambda n \mathcal{T}} \sum_{h=1}^{h_p} \pi \left[ \frac{\Delta}{2} l_p^h \right]^2 \leq c_p W \mathcal{T}. \quad (7)$$

By Cauchy-Schwarz Inequality,

$$\left[ \sum_{p=1}^{c_p \lambda n \mathcal{T}} \sum_{h=1}^{h_p} l_p^h \right]^2 \leq \left[ \sum_{p=1}^{c_p \lambda n \mathcal{T}} \sum_{h=1}^{h_p} (l_p^h)^2 \right] \left[ \sum_{p=1}^{c_p \lambda n \mathcal{T}} \sum_{h=1}^{h_p} 1 \right]. \quad (8)$$

Further, since there are at most  $n$  simultaneous transmissions at any given time in the network, we have

$$\sum_{p=1}^{c_p \lambda n \mathcal{T}} h_p \leq c_p W \mathcal{T} n. \quad (9)$$

Therefore,

$$\begin{aligned} \frac{16c_p W \mathcal{T}}{\pi \Delta^2} &\geq \sum_{p=1}^{c_p \lambda n \mathcal{T}} \sum_{h=1}^{h_p} (l_p^h)^2 \quad (\text{using (7)}) \\ &\geq \frac{\left[ \sum_{p=1}^{c_p \lambda n \mathcal{T}} \sum_{h=1}^{h_p} l_p^h \right]^2}{\left[ \sum_{p=1}^{c_p \lambda n \mathcal{T}} h_p \right]} \quad (\text{using (8)}) \\ &\geq \frac{(c_p \lambda n \mathcal{T} c)^2}{c_p W \mathcal{T} n} \quad (\text{using (6) and (9)}). \end{aligned}$$

Hence,

$$\lambda \leq \sqrt{\frac{16W^2}{\pi \Delta^2 c^2}} \frac{1}{\sqrt{n}}.$$

We are now ready to prove the main result of this section. We first define a general class of scheduling policies that we plan to study. Note that at each time instant and for each packet  $p$  that has not been delivered to its destination node yet, a scheduling policy essentially needs to make the following two types of decisions:

- *Replication:* The scheduler needs to decide whether to replicate the packet  $p$  to other relay nodes that do not have the packet yet. If yes, the scheduler needs to decide how to schedule radio transmissions to forward the packet  $p$  to these new relay nodes. Note that by *replication* we mean packet duplication; i.e., creating redundant copies of the packet. This is different from *capture* (to be defined next) where the number of copies of the packet does not increase.
- *Capture:* The scheduler needs to decide whether to deliver the packet  $p$  to the destination node immediately,

<sup>2</sup>A similar observation is used in [1] except that they take a receiver point of view.

possibly using multi-hop transmission. If yes, the scheduler needs to choose one relay node (possibly the source) that has a copy of packet  $p$  and schedule radio transmissions to forward the packet to the destination node. When this happens successfully, we say that the chosen relay node has successfully *captured* the destination node of packet  $p$ , or a successful capture has occurred for the packet  $p$ .

*Remark 4:* Although our model does allow for other less intuitive alternatives, in a typical scheduling scheme a successful *capture* usually occurs when a relay node holding the packet moves within a small neighborhood around the destination node, so that fewer resources are needed to forward the packet to the destination node. For example, a relay node could enter a disk of a certain radius around the destination node, or a relay node could enter the same cell as the destination node. We call such an area the *capture neighborhood*. The purpose of *replication* is to reduce the time before a successful *capture* occurs. With more nodes holding the packet  $p$ , the likelihood of one of them capturing the destination node sooner is higher.

In this paper, we restrict our study to the class of scheduling schemes that satisfy the following assumption:

**Assumption A:**

- Only the source of a packet is allowed to replicate the packet. That is, relay nodes holding a packet are *not* allowed to replicate it further.

*Remark 5:* Note that almost all scheduling schemes that have been proposed in the literature satisfy Assumption A [2]–[8].

It is worthwhile to elaborate on Assumption A, since it may seem restrictive at first sight. First, observe that the notions of *replication* and *relaying* are different, even though both involve forwarding packets to other relay nodes. For example, when node  $i$  decides to *replicate* the packet  $p$  to node  $j$ , node  $i$  can either transmit the packet directly to node  $j$ , or use multi-hop relaying; i.e., node  $i$  can forward the packet to another node  $k$ , and let node  $k$  *forward* the packet to node  $j$ . (Node  $k$  may also keep a copy of the packet  $p$ , in which case, both nodes  $k$  and  $j$  are considered to receive the packet due to *same* replication decision initiated by node  $i$ .) In this example, although both nodes  $i$  and  $k$  forward the packet  $p$  to other nodes, their roles are different. Node  $i$  is the one who *initiates* the replication, while node  $k$  is just *passively following* the instruction of node  $i$  to *relay* the packet to node  $j$ . Thus, we see that Assumption A only prohibits relay nodes from *initiating* a replication. In particular, multi-hop relaying is still allowed under Assumption A. (Multi-hop relaying is also allowed for the relay-to-destination communication, i.e., capture.)

If we attempt to develop *distributed* scheduling schemes, where nodes make replication decisions and capture decisions without any knowledge of the decisions at other nodes, then restricting the replication decisions to the source node is a natural way to *control the number of copies of a packet in the system*. Note that excessive redundancy can reduce the system throughput substantially. The source node of a packet  $p$  is in the best position to control both the total number of replications for the packet and the number of relay nodes

getting the packet in each replication. If the relay nodes were allowed to replicate, then additional cooperation among the relay nodes would likely be required (see, for example, the scheme in [13], where the relay nodes know the location of the static destination node, and also have some knowledge of the future direction of other nodes' movement, based on which they can *cooperate* to make selective and more efficient replication toward the destination node) in order to limit the number of replicas of a packet.

We can prove the following main result:

*Proposition 1:* Let  $\bar{D}$  denote the expected delay averaged over all packets and all source-destination pairs, and let  $\lambda$  denote the throughput of each source-destination pair. For any scheduling scheme that satisfies Assumption A, if

$$\bar{D} \leq O(n^\alpha/\sigma_n^2), \alpha < 0,$$

then

$$\lambda \leq O(1/\sqrt{n}).$$

*Proof:* Consider squares A and B of length  $1/4$ , centered at  $(-1/4, 1/4)$  and  $(1/4, -1/4)$ , respectively (see Fig. 3). Since the packet arrivals are independent of the positions of mobile nodes, there will be a constant fraction  $f_0$  of packets that have their source nodes in square A and destination nodes in square B, at the time of arrival. (If the stationary distribution of nodes positions is uniform, then  $f_0 = (\frac{1}{4})^4 = 1/256$ . Otherwise,  $f_0$  is still a positive constant independent of  $n$ .) Let  $\Phi_{AB}$  denote this set of packets. In order to ensure that  $\bar{D} \leq O(n^\alpha/\sigma_n^2)$ , the delay for packets in  $\Phi_{AB}$  has to be  $O(n^\alpha/\sigma_n^2)$ . Precisely, since  $\bar{D} \leq O(n^\alpha/\sigma_n^2)$ , there exists some  $N_0 > 0$  and  $c_1 > 0$ , such that

$$\bar{D} \leq c_1 n^\alpha / \sigma_n^2, \text{ when } n \geq N_0. \quad (10)$$

Therefore, the delay of at least half of the packets in  $\Phi_{AB}$  must be no greater than

$$t_n = \frac{2c_1 n^\alpha}{f_0 \sigma_n^2}.$$

(Otherwise, the delay of the other half of the packets in  $\Phi_{AB}$  must be greater than  $t_n$ . Because this other half contributes to at least  $f_0/2$  fraction of all packets, the condition (10) will be violated.) Let  $\Phi_{AB}^0$  denote the set of packets in  $\Phi_{AB}$  whose delay is no greater than  $t_n$ . Consider an arbitrary packet  $p$  which is in  $\Phi_{AB}^0$ . Let  $S_p$  and  $D_p$  denote its source node and destination node, respectively. Fig. 4 shows a typical packet delivery. The source nodes  $S_p$  moves from position  $S_0$  to  $U_1$ , and replicates the packet  $p$  to a relay node, say  $r_1$ , at position  $V_1$ , possibly using multi-hop transmission. The node  $r_1$  then moves independently of  $S_p$ . The source node moves on to position  $U_2$ , where it replicates the packet  $p$  to one more relay node, say  $r_2$ , positioned at  $V_2$ , and so on. It is also possible to replicate the packet to more than one relay node at the same time (for example, we can take  $U_1 = U_2$  if the source node replicates the packet to  $r_1$  and  $r_2$  at the same time). At time  $t \leq t_n$ , a successful capture occurs, as one of the relay nodes holding the packet  $p$  (node  $r_2$  in the case shown in Fig. 4) decides to forward the packet to its destination node  $D_p$ , which has moved from its initial position  $D_0$  to the position  $D$ , at time  $t$ . Let  $k_n$  denote the total number of

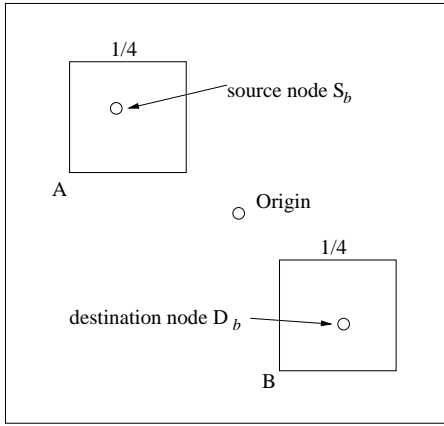


Fig. 3. There exists a constant fraction of packets that originate from nodes in A and are destined to nodes in B.

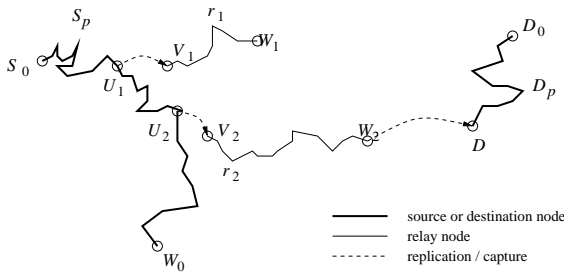


Fig. 4. How a typical packet  $p$  is delivered.

relay nodes that hold packet  $p$  in this process, and let  $r_k$ , for  $k = 1, 2, \dots, k_n$ , denote the  $k$ -th relay node. Let  $U_k$  and  $V_k$  denote the position of the source node  $S_p$  and the position of the relay node  $r_k$ , respectively, at time of replication. Let  $W_k$  denote the position of the relay node  $r_k$  at the time of capture (see Fig. 5). Since the direct straight-line path is always the shortest path connecting any two points, we have, for any  $k$ ,

$$d(S_0, U_k) + d(U_k, V_k) + d(V_k, W_k) + d(W_k, D) + d(D_0, D) \geq d(S_0, D_0).$$

Hence,

$$\begin{aligned} & d(U_k, V_k) + d(W_k, D) \\ & \geq d(S_0, D_0) - d(D_0, D) - d(S_0, U_k) - d(V_k, W_k). \end{aligned} \quad (11)$$

Since  $S_0$  and  $D_0$  are in the squares A and B, respectively,

$$d(S_0, D_0) \geq \frac{\sqrt{2}}{4}.$$

Further, each of the terms  $d(D_0, D)$ ,  $d(S_0, U_k)$ , and  $d(V_k, W_k)$ , corresponds to the movement of a different node. There are at most  $n$  nodes involved in this process. By setting  $c_n = 1/\log n$  in Corollary 1, we can see that, with probability approaching 1 as  $n \rightarrow \infty$ , all of the last three terms in (11) are no greater than  $\sqrt{2}/\log n$ , for all  $k$ . Therefore,

$$d(U_k, V_k) + d(W_k, D) \geq \frac{\sqrt{2}}{4} - 3\frac{\sqrt{2}}{\log n} \geq 1/4$$

for large enough  $n$ . Finally, let  $W_0$  denote the position of the

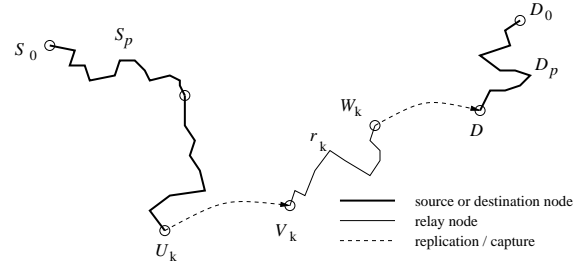


Fig. 5. The relay node  $r_k$ .

source node at time  $t$ . Then using a similar argument,

$$d(W_0, D) \geq d(S_0, D_0) - d(D, D_0) - d(S_0, W_0) \geq 1/4.$$

This shows that for each packet  $p$  in  $\Phi_{AB}^0$ , the total distance that the packet  $p$  has to be relayed is at least  $1/4$ . Since  $\Phi_{AB}^0$  contributes to at least  $f_0/2$  fraction of all packets, on an average each packet must be relayed over a distance no less than  $f_0/8 > 0$ . Hence, by Lemma 5, the per-node throughput  $\lambda$  must be no larger than  $O(1/\sqrt{n})$ . ■

*Remark 6:* For the ease of exposition, we have shown the above results for Brownian motion on a plane. However, it is not difficult to see that the argument in Proposition 1 applies to Brownian motion on a unit sphere as well. In particular, in Lemma 4, if we choose  $c_n = c/\log n$ , the size of the square  $\mathbf{SQ}(c_n)$  diminishes to zero as  $n \rightarrow \infty$ . Hence, the difference between such a square on a plane and that on a unit sphere vanishes. Therefore, both Corollary 1 and Proposition 1 hold for Brownian motion on a unit sphere as well.

**The Degenerate Tradeoff:** Proposition 1 shows that the *capacity-delay trade-off under the Brownian motion model is degenerate*: For delay less than  $O(n^\alpha/\sigma_n^2)$ ,  $\alpha < 0$ , the per-node throughput is at most  $O(1/\sqrt{n})$ . Since a per-node throughput of  $\Theta(1/\sqrt{n})$  can be achieved under the static settings, using multi-hop relaying [1], our result shows that whenever the delay is constrained to be less than  $O(n^\alpha/\sigma_n^2)$  for some  $\alpha < 1$ , Brownian mobility does not result in any improvement of the throughput. Further, since the packet transmissions are usually carried out at a much faster time-scale than the node mobility, one could view the delay under the multi-hop scheduling (see [1]) as being *almost zero*. Earlier studies have shown that it is possible to achieve  $\Theta(1)$  per-node throughput at roughly  $\Theta(1/\sigma_n^2)$  delay under the Brownian motion model. Obviously,  $\Theta(1)$  is an upper bound on the per-node capacity (under our network model). Hence, if we ignore the logarithmic terms, the capacity-delay trade-off under the Brownian motion model degenerates into two points: one can either achieve a per-node throughput of  $\Theta(1/\sqrt{n})$  at almost no delay, or a per-node throughput of  $\Theta(1)$  at roughly  $\Theta(1/\sigma_n^2)$  delay, but nothing in between! Finally, although Proposition 1 is shown under the Brownian motion model, it is not difficult to see that the result also applies to the Markovian mobility model in [4]. This is because as  $n \rightarrow \infty$ , the difference between these mobility models vanishes.

The result of Proposition 1 is in sharp contrast to the results reported in existing works [3], [6], where it is claimed that certain schemes can provide a smooth trade-off between

the capacity and delay. Since the schemes in [3], [6] satisfy Assumption A, it is clear that they cannot provide a smooth delay-capacity trade-off.

## V. DELAY UNDER GENERALIZED TWO-HOP RELAYING SCHEMES

In Section IV, we have established the fundamental delay-capacity trade-off under the Brownian motion model for a wide class of scheduling schemes. We have shown that in case of a scheduling scheme that satisfies Assumption A, in order for the per-node throughput to be  $\Omega(1/\sqrt{n})$ , the delay must be  $\Omega(n^\alpha/\sigma_n^2)$  for all  $\alpha < 0$ . In this section, we study the delay performance of a more restricted set of scheduling schemes. Our interest in this class of schemes stems from the fact that they have been used in the earlier studies for achieving  $\Omega(1/\sqrt{n})$  per-node throughput, under various mobility models. We now investigate their delay performance under the Brownian motion model.

These schemes are generalizations of the 2-hop relaying scheme of Grossglauser and Tse [2]. Hence, we refer to them as *generalized 2-hop relaying schemes*. Compared with the more general class of schemes that we considered in Section IV, these schemes have one additional restriction: For each packet  $p$ , the source node is only allowed to replicate the packet  $p$  to *one* relay node (denoted by  $R(p)$ ). Other than this restriction, the generalized 2-hop relaying schemes still have substantial flexibility in scheduling packet transmissions. For example, in the *replication phase*, the scheduler still decides when to replicate the packet, and how (e.g., which relay node to replicate the packet  $p$  to, and how to schedule the packet transmissions from the source node to the chosen relay node  $R(p)$ , possibly using multi-hop transmissions). Similarly, in the *capture phase*, the scheduler decides when and how to relay the packet  $p$  to the destination node, from either the source node or the chosen relay node  $R(p)$ , possibly using multi-hop transmissions.

To ensure that fewer radio resources are consumed, the replication phase (correspondingly, capture phase) typically occurs when the chosen relay node is within a small neighborhood around the source node (correspondingly, destination node). For example, a relay node could either enter a disk of a certain radius around the source node (or destination node), or a relay node could enter the same cell as the source node (or destination node) in case the network is divided into cells. We call such an area around the source node or the destination node as the *replication neighborhood* or the *capture neighborhood*, respectively. We further assume that the replication neighborhood and the capture neighborhood are both contained in disks of radius  $a_n$  centered at the source node and the destination node, respectively. Again, to ensure that fewer radio resources are consumed,  $a_n$  would typically be  $o(1)$ .

*Remark 7:* Note that Scheme 2 and Scheme 3(b) in [3] are both special cases of the generalized 2-hop relaying schemes that we consider in this section.

We now give a lower bound on the average packet delay under the generalized 2-hop relaying schemes.

*Proposition 2:* If the replication neighborhood and the capture neighborhood under a generalized 2-hop relaying scheme can be contained inside a disk of radius  $a_n$  around the source node and destination node, respectively, and  $a_n = o(1)$ , then the average packet delay under the given scheduling scheme must be  $\Omega(\log(1/a_n)/\sigma_n^2)$ .

*Proof:* When  $a_n = o(1)$ , most packets will have to be delivered through a relay node. Consider such a random packet that arrives at the source node. Its delay must be no less than the time that it takes for the relay node to move from somewhere within distance  $a_n$  around the source node, to somewhere within distance  $a_n$  around the destination node. Since the packet arrival processes are independent of the node mobility processes, the source node and destination node will be distributed uniformly inside the network, at the time of packet arrival. Therefore, the delay for the packet will be no less than the time that it takes for two nodes placed uniformly inside the network, to come within a distance of  $2a_n$  from each other. Therefore, in view of Lemma 2, the result follows. ■

*Remark 8:* Note that Proposition 2 holds even if the replication neighborhood and capture neighborhood are different in shape or size. Further, if  $a_n = O(n^\alpha)$  for some  $\alpha < 0$ , then the delay under the generalized 2-hop relaying scheme is

$$\Omega(\log n/\sigma_n^2). \quad (12)$$

The above result provides a lower bound on the average packet delay under any generalized 2-hop relaying scheme. We have not provided the analysis for the upper bound on delay. Using the methodology in [3] and making some technical assumptions, it can be shown that the delay is in fact  $\Theta(\log n/\sigma_n^2)$ .

## VI. DISCUSSION OF RELATED WORKS

In this paper, we have investigated the capacity-delay trade-off for mobile ad hoc networks under the Brownian motion model. We showed that the capacity-delay trade-off under the Brownian motion model is degenerate. For delays smaller than  $O(n^\alpha/\sigma_n^2)$  for some  $\alpha < 0$ , the per-node throughput is at most  $O(1/\sqrt{n})$ , while a per-node throughput of  $\Theta(1)$  can be obtained while incurring a delay of  $\Theta(\log n/\sigma_n^2)$ .

We would like to point out that the results of this paper are asymptotic in nature, and having *virtually* no tradeoff between the delay and capacity, does not rule out the possibility of some limited trade-off between the delay and capacity. In case of finite systems, even the constants before the order results can be significant. Moreover, our results do not rule out the possibility of having an average packet delay of  $\Theta(n/k_n)$ , where  $k_n = o(n^\alpha)$  for all  $\alpha > 0$ , and a capacity of  $\omega(1/\sqrt{n})$ .

Earlier work on delay-capacity trade-off under the Brownian motion model can be found in [3], [6]. Although, both these works consider the Brownian motion model, the delay-capacity trade-offs reported in these works differ substantially from ours. In particular, we showed that the achievable delay-capacity trade-off under the Brownian motion model is degenerate, while both [3] and [6] report smooth trade-offs. We now briefly point out the reasons for this discrepancy.



We first look at the results in [3]. The errors in [3] follow from an incorrect embedding of the Brownian motion model to a random walk model. More precisely, the authors divide the unit torus into  $\frac{1}{\sqrt{a(n)}} \times \frac{1}{\sqrt{a(n)}}$  cells of equal size, where  $1/n \leq a(n) < 1$ . They then consider a random walk model on this grid with a jump time of  $\Theta(\sqrt{a(n)/v(n)})$ , where  $v(n)$  is defined in [3] as the speed of nodes. However, from Remark 2, it follows that if the underlying mobility model is a Brownian motion model on a torus then the jump times should be proportional to the area  $a(n)$  of the cells, rather than  $\sqrt{a(n)}$ . Furthermore, there is also incorrect, by a factor of  $\log n$ , estimation of the delay in [3]. A detailed discussion of these issues can be found in our technical report [14].

This paper also corrects the previously reported results in [6]. In [6], a smooth delay-capacity trade-off is reported as the number of relay nodes per packet is varied. The derivation in [6] assumes that the paths of all mobile relays are independent of each other. This assumption, however, does not hold since all relay nodes receive the packet from the *same* source node, and hence the starting points of their paths are highly correlated to each other. Indeed, by Lemma 4, even if the source node replicates the packet to  $\Theta(n)$  relay nodes, the delay is not reduced much, i.e., it is still close to  $\Theta(1/\sigma_n^2)$ . This is in contrast to the random way-point mobility model (which is also studied in [6]), where the correlation between the paths of various relay nodes holding the packet dies out in a very short time, allowing a smooth trade-off between the delay and capacity [5].

## VII. CONCLUDING REMARKS

In this paper, we studied the fundamental trade-off between the delay and capacity under the Brownian motion model. We have shown that the capacity-delay trade-off under the Brownian motion model is *degenerate*: one can achieve a per-node throughput of  $\Theta(1)$  with  $\Omega(\log n/\sigma_n^2)$  delay (using 2-hop relaying), but even when the delay is constrained to be  $O(n^\alpha/\sigma_n^2)$  for some  $\alpha < 0$ , one can only achieve a per-node throughput of  $\Theta(1/\sqrt{n})$ , which is the same as the achievable throughput under the static setting.

This paper, along with the related results reported in the past [3]–[8], provides a much better understanding of the relationship between the delay and capacity in mobile ad hoc networks. Interestingly, it turns out the delay-capacity trade-off critically depends on the underlying mobility model. As we have shown in this paper, the trade-off is *degenerate* under the Brownian motion model. On the other hand, under the *i.i.d.* mobility model, we have shown in [8] that the trade-off between the per-node throughput  $\lambda$  and delay  $D$  satisfies:

$$\lambda \leq \Theta\left(\sqrt[3]{\frac{D}{n}} \log n\right)$$

for  $\Theta(1) \leq D \leq \Theta(n)$ , and the above upper bound is achievable up to a logarithmic factor, which indicates that the bound is tight. (Note that under the *i.i.d.* mobility model, each node randomly picks its position at each time slot, independently of either the other nodes' positions or its own position in the past.) Finally, under the random way-point

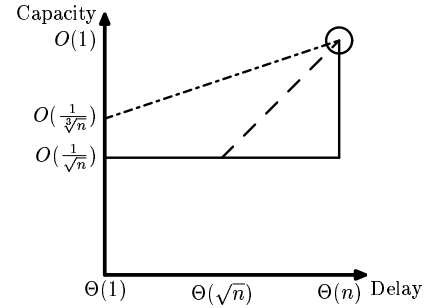


Fig. 6. The delay-capacity trade-offs under the Brownian motion model (the solid line), the random way-point mobility model (the dashed line), and the *i.i.d.* mobility model (the dash-dotted line). We have chosen  $v_n = \sigma_n = 1/\sqrt{n}$  and ignored all logarithmic terms in the figure.

mobility model, we have shown in [5], [6] that the following delay-capacity trade-off can be achieved:

$$\lambda = \Theta\left(\frac{Dv_n}{\sqrt{n}}\right)$$

for  $\Theta(1/v_n) \leq D \leq \Theta(\sqrt{n}/v_n)$ , where  $v_n$  is the speed of the nodes. (Note that under the random way-point mobility model, each node picks a random destination and moves toward it with speed  $v_n$ . Once the node reaches that destination, it then picks another destination randomly and moves toward it, and the process continues.) In Fig. 6, we illustrate the difference in the three delay-capacity trade-offs we have obtained. In this figure, we have chosen  $v_n = \sigma_n = 1/\sqrt{n}$ . The reason for such choice of  $v_n$  and  $\sigma_n$  is to ensure that the *contact time* (i.e., the time for two nodes to remain neighbors of each other) is  $\Theta(1)$  under all three mobility models. As we can see, a smooth trade-off exists for any value of delay for the *i.i.d.* mobility model, while a smooth trade-off only exists for delay between  $\Theta(\sqrt{n})$  and  $\Theta(n)$  under the random way-point mobility model, and the trade-off degenerates to only two points under the Brownian motion model.

Looking at these results, it is natural to ask: What will the delay-capacity trade-off be for a *real* mobile wireless network? Will the trade-off in real networks be one of these three types? Or will it be a combination of these three? A closely related question is whether these trade-offs (along with their respective mobility models) represent three *distinct* cases, or they are part of a continuous range of delay-capacity trade-offs. These and other issues are addressed elsewhere [15].

## REFERENCES

- [1] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. on Information Theory*, vol. IT-46, no. 2, pp. 388–404, March 2000.
- [2] M. Grossglauser and D. N. C. Tse, "Mobility Increases the Capacity of Ad-hoc Wireless Networks," in *IEEE INFOCOM*, 2001, pp. 1360–1369.
- [3] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-Delay Trade-off in Wireless Networks," in *IEEE INFOCOM*, 2004.
- [4] M. Neely and E. Modiano, "Capacity and Delay Tradeoffs for Ad-Hoc Mobile Networks," to appear in *IEEE Trans. on Information Theory*, 2005.
- [5] G. Sharma and R. Mazumdar, "Delay and Capacity Trade-off in Wireless Ad Hoc Networks with Random Way-point Mobility," Preprint, School of ECE, Purdue University, Available at <http://ece.purdue.edu/~gsharma>, 2005.

- [6] —, “On Achievable Delay/Capacity Trade-offs in Mobile Ad Hoc Networks,” in *Workshop on Modeling and Optimization in Ad Hoc Mobile Networks*, 2004.
- [7] S. Toumpis and A. Goldsmith, “Large Wireless Networks under Fading, Mobility, and Delay Constraints,” in *IEEE INFOCOM*, 2004.
- [8] X. Lin and N. B. Shroff, “The Fundamental Capacity-Delay Tradeoff in Large Mobile Ad Hoc Networks,” in *Third Annual Mediterranean Ad Hoc Networking Workshop*, 2004.
- [9] M. Franceschetti, O. Dousse, D. Tse, and P. Thiran, “On The Throughput Capacity of Random Wireless Networks,” Preprint, Dept. of ECE, University of California, San Diego. Available at: <http://fbeece.ucsd.edu/~massimo/>, 2005.
- [10] A. Agarwal and P. R. Kumar, “Capacity Bounds for Ad-Hoc and Hybrid Wireless Networks,” *ACM SIGCOMM Computer Communications Review, Special Issue on Science of Networking Design*, vol. 34, no. 3, pp. 71–81, July 2004.
- [11] D. Brillinger, “A Particle Migrating Randomly on a Sphere,” *J. Theoret. Probab.*, vol. 10, no. 2, pp. 429–443, April 1997.
- [12] R. Durrett, *Probability : Theory and Examples*, 2nd ed. Belmont, CA: Duxbury Press, 1996.
- [13] N. Bansal and Z. Liu, “Capacity, Delay and Mobility in Wireless Ad-Hoc Networks,” in *IEEE INFOCOM*, April 2003.
- [14] X. Lin, G. Sharma, R. Mazumdar, and N. Shroff, “Degenerate Delay-Capacity Trade-offs in Ad Hoc Networks with Brownian Mobility,” Technical Report, School of ECE, Purdue University, Available at <http://ece.purdue.edu/~gsharma>, 2005.
- [15] G. Sharma, R. Mazumdar, and N. Shroff, “Delay and Capacity Trade-offs in Mobile Ad Hoc Networks: A Global Perspective,” *IEEE INFOCOM 2006*, 2005.