

Delay and Capacity Trade-offs in Mobile Ad Hoc Networks: A Global Perspective

Gaurav Sharma* , Ravi Mazumdar†, Ness Shroff*

* School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47907, USA

Email: {gsharma,shroff}@ecn.purdue.edu

† Department of Electrical and Computer Engineering
University of Waterloo

Waterloo, ON N2L 3G1, Canada

Email: mazum@ece.uwaterloo.ca

Abstract—Since the original work of Grossglauser and Tse, which showed that the mobility can increase the capacity of an ad hoc network, there has been a lot of interest in characterizing the delay-capacity relationship in ad hoc networks. Various mobility models have been studied in the literature, and the delay-capacity relationships under those models have been characterized. The results indicate that there are trade-offs between the delay and the capacity, and that the nature of these trade-offs is strongly influenced by the choice of the mobility model. Some questions that arise are: (i) How representative are these mobility models studied in the literature? (ii) Can the delay-capacity relationship be significantly different under some other “reasonable” mobility model? (iii) What would the delay-capacity trade-off in a real network be like? In this paper, we address these questions. In particular, we analyze, among others, some of the mobility models that have been used in the recent related works, under a unified framework. We relate the nature of the delay-capacity trade-off to the nature of the node motion, thereby providing a better understanding of the delay-capacity relationship in ad hoc networks than earlier works.

I. INTRODUCTION

In this paper, we study the delay and capacity trade-offs in mobile ad hoc networks. We study the network layer notion of capacity, first introduced by Gupta and Kumar [1]. In [1], it was shown that the per-node capacity of a random wireless network with static nodes scales as $\Theta\left(\frac{1}{\sqrt{n \log n}}\right)$, provided the same power level is used by all the nodes in the network. Recent works [2], [3] have shown that a per-node capacity of $\Theta(1/\sqrt{n})$ can be achieved when the nodes are allowed to exercise power control.

Grossglauser and Tse [4] have shown that significant gains in the per-node capacity can be obtained by exploiting the mobility of the nodes in a mobile ad hoc network. In particular, they proposed a 2-hop relaying scheme, and showed that their scheme can achieve a constant per-node capacity, which does not vanish as the number of nodes grows arbitrarily large. The delay related issues were not addressed in [4].

Since both capacity as well as delay are important from an application point of view, a significant effort has recently been devoted within the networking research community to understand the delay-capacity relationship in ad hoc networks.

Bansal and Liu [5] were among the first to study the delay-capacity relationship in wireless networks. They considered a wireless network consisting of static sender-destination pairs and mobile relays, and proposed a geographic routing scheme that achieves a near optimal capacity and studied its delay performance. Perevalov and Blum [6] studied the delay limited capacity of mobile ad hoc networks. Their work was motivated by the *diversity coding* approach given in [7].

Recent works of Neely and Modiano [8], ElGamal et al. [9], Toupis and Goldsmith [10], Sharma and Mazumdar [11], [12], Lin and Shroff [13], and Lin et al. [14], have all studied the delay-capacity trade-offs in mobile ad hoc networks. The mobility models and the network settings studied in these works differ considerably, and so do the delay-capacity trade-offs that are reported. The mobility models studied in the literature include the i.i.d. mobility model [8], [10], [13]; the random way-point mobility model [11], [12]; the Brownian mobility model [11], [14], [9]; and the random walk mobility model [9], [15].

Since the network settings considered in these works are quite different, it is difficult to single out the impact the nature of the node mobility has on the delay-capacity trade-off. For example, both [8] and [13] study the i.i.d. mobility model, but the trade-offs reported in these works are quite different. In particular, in [8], the authors consider a cell partitioned network setting and show that the trade-off: $\lambda \leq \Theta\left(\frac{D}{n}\right)$, where D is the average packet delay and λ is the average per-node throughput, is both necessary as well as sufficient under their setting. Under a less restrictive network setting in [13], the delay-capacity trade-off is shown to be: $\lambda^3 \leq \Theta\left(\frac{D \log^3 n}{n}\right)$. Thus, we see that the delay-capacity trade-off depends not only on the nature of the node mobility, but also on the network setting.

The main insight we provide in this paper is that there is a critical value of delay (henceforth denoted by *critical delay*), below which, the node mobility cannot be exploited for improving the capacity. We also show that the critical delay depends on the nature of the node mobility, but not so much

on the network setting. In terms of the notion of critical delay (see section III for a rigorous definition), some recent results in the literature can be summarized as follows:

- The critical delay under the Brownian motion model and random walk model is roughly $\Theta(n)^*$ (see [14], [15]).
- The critical delay under the random way-point mobility model is roughly $\Theta(\sqrt{n})$ (see [12]).
- The critical delay under the i.i.d. mobility model is $\Theta(1)$ (see [13]).

Observing the above results, it is natural to ask: (i) How representative are the above mentioned mobility models? (ii) Can the critical delay and, in general, the delay-capacity relationship be significantly different under some other mobility model? (iii) How does the critical delay scale in a real network? This paper makes the following contributions toward answering these fundamental questions:

- We propose and study a family of *hybrid random walk models*, and show that they exhibit a continuous range of critical delays in-between those of the i.i.d. mobility model and the random walk mobility model. In particular, for every β between 0 and 1, there exists a mobility model in the family of *hybrid random walk models* for which the critical delay is roughly $\Theta(n^\beta)$. As expected, $\beta = 1$ corresponds to random walk mobility model, and $\beta = 0$ corresponds to the i.i.d. mobility model.
- We propose and study a family of *discretized random direction models* exhibiting a continuous range of critical delays in-between those of the random way-point mobility model and the Brownian mobility model. In particular, for every α between 0 and $1/2$, there exists a mobility model in the class of *discretized random direction models* for which the critical delay is roughly $\Theta(n^{1/2+\alpha})$. Note that $\alpha = 0$ corresponds to the random way-point mobility model, and $\alpha = 1/2$ corresponds to the Brownian mobility model. These models approximate the motion of the nodes under the commonly used random direction models in the literature (see, for example, [16]), and are simpler to analyze.

An interesting feature of the above classes of mobility models is that all of them incur a delay of roughly $\Theta(n)$ under the 2-hop relaying scheme; which is in line with the other mobility models considered in the literature. Our results therefore show that the mobility models considered in the literature are in some sense extreme: they either exhibit the smallest critical delays (i.i.d. mobility model and random way-point mobility model) or the largest critical delays (Brownian motion model and random walk mobility model), among the mobility models in their respective classes.

The rest of the paper is organized as follows. We introduce the *hybrid random walk models* and *discretized random direction models*, and discuss our network and transmission model in the next section. We define the notions of critical delay and

2-hop delay in section III. We study the critical delay and 2-hop delay under the hybrid random walk models in section IV, and under the discretized random direction models in section V. A discussion on the implications of our results for large mobile ad hoc networks is provided in section VI. Finally, we end this paper with some concluding remarks in section VII.

II. THE MODEL

A. The Network and Transmission Model

We consider an ad hoc network consisting of n mobile nodes, distributed uniformly on a unit square. We consider a homogeneous scenario in which each node generates traffic at the same rate. Further, we assume that each node, say node i , generates traffic for exactly one other node, say node $d(i)$, and that the mapping $i \mapsto d(i)$ is bijective. We also assume that the packet arrival process at each node is independent of the node mobility process.

The communication between any *source-destination* pair can possibly be carried out via multiple other nodes, acting as relays. That is, a *source* node can, if possible, send a packet directly to its *destination* node; or, the source node can forward the packet to one or more *relay* nodes; the relay nodes can also forward the packet to other relay nodes; and finally, a relay node or the source node itself can deliver the packet to its destination node.

For simplicity, we assume that the success or failure of a transmission between a pair of nodes is governed by 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 node i , for $i = 1 \dots n$, at time t . Under the protocol model, node i can communicate directly with node j at a 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 Δ is some positive number; and $d(x, y)$ denotes the Euclidean distance between points $x, y \in \mathbb{R}^2$. Note that for a packet to be successfully received by node j , the above interference constraint must be satisfied over the entire duration of the packet transmission from node i to node j .

We use the following definition of throughput: Let $\lambda_i(t)$ be the total number of bits delivered end-to-end for source-destination pair i upto time t , then the throughput $\lambda(n)$ of the system is given by

$$\lambda(n) = \liminf_{t \rightarrow \infty} \frac{\sum_{i=1}^n \lambda_i(t)}{t}.$$

Next, we describe our mobility models, starting with the hybrid random walk models.

B. Hybrid Random Walk Models

These models are parametrized by a single parameter β , which takes values between 0 and $1/2$. The unit square is divided into $n^{2\beta}$ squares of area $1/n^{2\beta}$ each (henceforth

*Note that in [14], the results are stated in terms of the variance parameter $\sigma^2(n)$. We have set $\sigma^2(n) = 1/n$ for the sake of easy comparison with the other results.

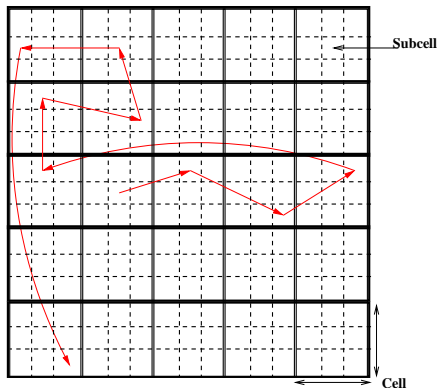


Fig. 1. The figure shows the division of the unit square into cells and subcells and also the motion of a node under the hybrid random walk model.

referred to as cells), resulting in a discrete torus of size $n^\beta \times n^\beta$. Each cell is then further divided into $n^{1-2\beta}$ square subcells of area $1/n$ each[†], as shown in Fig. 1. Time is divided into slots of equal duration. At each time slot a node is assumed to be in one of the subcells inside a cell. Initially, each node is equally likely to be in any of the n subcells, independent of the other nodes. At the beginning of a time-slot, a node jumps from its current subcell to one of the subcells in an adjacent cell, chosen in a uniformly random fashion. By adjacent cell we mean the following: Let $(i, j) : i, j = 0, 1, \dots, n^\beta - 1$, be a numbering of the cells of the 2-d torus, as shown in Fig. 2. The cells adjacent to cell (i, j) are the cells $(i+1, j)$, $(i-1, j)$, $(i, j+1)$, and $(i, j-1)$, where the addition and subtraction operations are performed modulo n^β . Note that for $\beta = 0$, the above mobility model is essentially the i.i.d. mobility model considered in [8], [13], [10]; and for $\beta = 1/2$, it is the same as the random walk model of [9].

Next, we describe the random direction models (see, for example, [16]).

C. Random Direction Models

These models are parametrized by a single parameter α , which takes values between 0 and $1/2$. The initial position of each node is assumed to be uniformly distributed inside the unit square. The motion of each node under these models is independent and identical to the other nodes. The motion of a node is divided into multiple trips. At the beginning of a trip, the node chooses a direction θ uniformly between $[0, 2\pi]$, and moves a distance of $n^{-\alpha}$ in that direction, with a speed v_n , and the process repeats itself. Note that we are assuming a complete warp-around of the unit square (see Fig. 3).

Remark 1. *In the rest of the paper, we consider $v_n = \Theta(1/\sqrt{n})$, as in [12]. This particular choice of node speed is motivated by the fact that we keep the network area fixed and let the number of nodes increase to infinity, which means that the average neighborhood size scales as $\Theta(1/\sqrt{n})$. In order*

[†]Throughout this paper we ignore the issues pertaining to $n^{1-2\beta}, n^{2\beta}$ not being perfect squares.

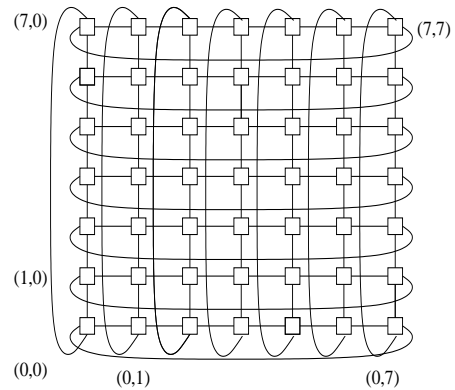


Fig. 2. A numbering of cells on a 2-d torus.

to account for this “shrinking” of the neighborhood size, we scale the node velocity as $\Theta(1/\sqrt{n})$. Note that, alternatively, one could scale the area of the network in proportion to n , while keeping the velocity of the nodes fixed.

Next, we describe the discretized random direction models, that we study in this paper. As discussed before, these models approximate the motion of the nodes under the random direction models, and are simpler to analyze.

D. Discretized Random Direction Models

These models, like the random direction models, are parametrized by a single parameter α , which takes values between 0 and $1/2$. The unit square is divided into $n^{2\alpha}$ squares of area $1/n^{2\alpha}$ each (henceforth referred to as cells), resulting in a discrete torus of size $n^\alpha \times n^\alpha$.

Time is divided into slots of equal duration. At the beginning of a time slot, each node jumps from its current cell to an adjacent cell, chosen uniformly from within the set of adjacent cells. The motion of a node during the time slot is as follows: The node chooses a start and end point uniformly from within the current cell. During the time slot, the node moves from the start point to the end point. In order to keep the duration of all time slots the same, the velocity of the node is chosen to be inversely proportional to the distance between the start and the end point.

In view of Remark 1, the duration of a time slot should be $\Theta(n^{1/2-\alpha})$. In order to be able to compare these models with the Brownian motion model of [11], [14], we consider $\sigma_n^2 = 1/n$, where σ_n^2 is the variance parameter of the Brownian motion model, as defined in [11], [14]. Note that with the above choice of σ_n^2 and v_n , each node moves an average distance of $\Theta(1/\sqrt{n})$ in unit time, under all these models. Also observe that under these settings, the discretized random direction model with $\alpha = 1/2$ degenerates into the random walk model, which is the discrete time version of the Brownian motion model.

III. PRELIMINARIES: CRITICAL DELAY, 2-HOP DELAY, SCHEDULING SCHEMES

In this section, we build the platform required for the study of the critical delay in the subsequent sections. We first start

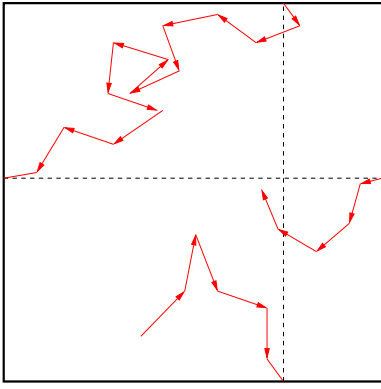


Fig. 3. The figure shows the motion of a node under the random direction model.

by providing a rigorous definition of the critical delay and motivating its study.

It is now well known that the per-node throughput capacity of static ad hoc networks scales as $O(1/\sqrt{n})$, with the capacity achieving scheme being the famous multihop relaying scheme. Grossglauser and Tse [4] showed that a constant throughput per-node can be achieved under any stationary and ergodic mobility model which preserves the uniform distribution of the nodes at all times. The delay performance of the capacity achieving 2-hop relaying scheme of [4] has been studied in many recent works [8], [9], [13], [11], [12], [15], and the 2-hop relaying scheme has been shown to incur an average delay of about $\Theta(n)$ under many different mobility models. Alternative schemes that provide a throughput in between that of multihop relaying and 2-hop relaying have also been studied in the literature, under various mobility models. The objective of these schemes is to provide better delay performance than the 2-hop relaying scheme by compromising the throughput capacity. An important quantity to study in this respect is what we call the critical delay, which we next define:

Definition 1. Let \mathcal{C} be the class of scheduling and relaying schemes under consideration, and let D_c, λ_c be the average delay and the per-node throughput, respectively, under scheme $c \in \mathcal{C}$. The critical delay under the class of schemes \mathcal{C} , denoted by $D_{\mathcal{C}}$, is the minimum average delay that must be tolerated under a given mobility model in order to achieve a per-node capacity of $\omega(1/\sqrt{n})$, that is,

$$D_{\mathcal{C}} = \inf_{\{c \in \mathcal{C} : \lambda_c = \omega(1/\sqrt{n})\}} D_c. \quad (2)$$

Remark 2. Note that the requirement of $\omega(1/\sqrt{n})$ capacity in the above definition is to ensure that the asymptotic throughput is above that of static ad hoc networks.

Remark 3. Note that each scheme c in the class of schemes \mathcal{C} is actually sequence of schemes c_n , where c_n is the scheme used with n nodes in the network. Similarly, λ_c and D_c are also sequences rather than numbers. The infimum in Eq.(2) should therefore be interpreted as a sequence of infimums, one for each n .

Remark 4. From the above definition, it is clear that the larger the class \mathcal{C} is, the smaller the critical delay would be. Ideally, one would like the class \mathcal{C} to include all possible scheduling schemes. In that case, if the delay that can be tolerated is smaller than the critical delay for a given mobility model, then one cannot exploit the node motion under that mobility model to increase the throughput capacity (in order sense) beyond its value under static conditions.

Henceforth, we will denote by 2-hop delay, the delay under the 2-hop relaying scheme. Observe that the trade-off between the delay and the capacity exists for delay values that are greater than the critical delay, and smaller than the 2-hop delay. Next, we introduce two key notions: *first hitting time* and the *first exit time*, which will be used to study the 2-hop delay and critical delay, respectively.

We start with some notation. Let S denote the unit square centered at the origin, and let x, y be two points inside S . We denote by $[x]$ the set of points $\{x-1, x, x+1\}$, and define

$$d([x], [y]) = \min_{x_i \in [x], y_j \in [y]} d(x_i, y_j).$$

We denote by $B(x, r)$ the set of points y inside S such that $d([x], [y]) \leq r$. Let X_i^t denote the position of the node i at time t , under some mobility model. Note that the index t could either be continuous or discrete, depending on the mobility model. We are now ready to define the notion of the first exit time:

Definition 2 (First Exit Time). Let $X_i^0 = x$. The first exit time of $B(x, r)$, denoted by τ_E^r , is given by:

$$\tau_E^r = \inf\{t \geq 0 : X_i^t \notin B(x, r)\}.$$

It should be clear from the above definition that the statistical properties of the exit time do not depend on the choice of the point x and node i . Also note that the notion of exit time is well studied in the mathematics literature, under a variety of contexts. Our interest in notion of exit time stems from the fact that it has a close connection with the critical delay, which will be exploited in the subsequent sections.

We will next define the first hitting time. In our case, we find it more convenient to define it in discrete time. In order to do so, we need some more notation. Let $X(t)$ be a Markov chain on state space S , with stationary distribution Π . We have the following definition:

Definition 3 (First Hitting Time). The first hitting time for the set of states A is given by:

$$\tau_H^A = \inf\{t \geq 0 : X(t) \in A\},$$

with $X(0)$ being distributed according to Π .

Again, we would like to point out that the notion of the first hitting time has been widely studied in the mathematics literature, under many different contexts. As will be discussed in subsequent sections, the first hitting time has a close connection with the 2-hop delay as well as the critical delay. This connection has been exploited in several recent works

for estimating the 2-hop delay under various mobility models (see, for example, [12], [14]).

Next, we recall the result concerning the first hitting time for a single state in case of a 2-d torus of size $\sqrt{n} \times \sqrt{n}$ (for a proof see, for example, [14]):

Lemma 1. *Let H denote the first hitting time for a single state on a 2-d torus of size $\sqrt{n} \times \sqrt{n}$, then $\mathbb{E}\{H\} = \Theta(n \log n)$.*

The above result will be used quite often during the analysis in subsequent sections.

We now define the class of scheduling schemes that we will study in this paper. The only restriction we place on our scheduling schemes is the following:

Assumption A:

- Only the source node is allowed to initiate a replication, that is, the relay nodes holding a packet are *not* allowed to initiate a replication.

Note that almost all scheduling schemes that have been considered in the literature satisfy Assumption A. To understand this, we must elaborate on the notion of replication. By *replication* we mean packet duplication, that is, creating redundant copies of a packet. This is different from multihop relaying where a single copy of a packet exists in the network. It is important to note that the notions of *replication* and *relaying* are different, even though both involve forwarding packets to other relay nodes. To understand this, suppose node i decides to *replicate* a packet at node j ; then node i can either transmit the packet directly to node j ; or use multihop relaying, where the packet reaches node j through multiple intermediate nodes. Note that, if asked by node i , the intermediate nodes may also keep a copy of the packet with them, and in that case, all of them are considered to receive the packet due to the *same* replication decision initiated by node i . In this example, although both node i and the other intermediate nodes forward the packet to other nodes, their roles are different. Node i is the one that *initiates* the replication, while the intermediate nodes *passively follow* the instructions of node i . Thus, we see that Assumption A only prohibits relay nodes to *initiate* replication and, in particular, multihop relaying is allowed under Assumption A.

We also allow *immediate capture* of the destination using multihop relaying or long range wireless broadcast, at any time during the replication process. Note that by capture we mean the successful delivery of a packet to its destination node. Although we allow for other less intuitive alternatives, in a typical scheduling scheme a successful *capture* usually occurs when a relay node holding the packet comes within a small area around the destination node, so that fewer resources are needed to forward the packet to the destination. For example, a relay node could enter a disk of a certain radius around the destination, or a relay node could enter the same cell as the destination. We call such an area the *capture neighborhood*. The purpose of *replication* is to reduce the time before a successful *capture* occurs (since with more nodes holding the packet, the likelihood of one of them capturing the destination sooner is higher).

We use the word “immediate” to emphasize that capture can be carried out (using multihop relaying or long range wireless broadcast, if required) at a much faster time-scale than the node mobility, and the same is true of the replication process as well. The reason for this is that the packet transmission is usually carried out at a much faster time scale than the node mobility, and as a result the change in the position of a node during a packet transmission is often negligible.

Note that almost all scheduling schemes that have been studied in the literature satisfy Assumption A [4], [9], [8], [12], [11], [10], [13], [15]. The reason for this, we believe, is that in a distributed system, where nodes make replication decisions and capture decisions without any knowledge of the decisions at the other nodes, restricting the replication decisions to the source node is a natural way to *control the number of copies of a packet*. Note that higher redundancy implies smaller throughput. The source node of a packet is in the best position to control both the total number of replications of the packet and the number of relay nodes getting a copy of the packet with each replication. If the relay nodes were allowed to initiate replication, then additional cooperation among the relay nodes would likely be required in order to limit the number of replicas of a packet. An interesting example of this would be the scheme of Bansal and Liu [5], where the relay nodes know the location of the static destination node, and also have some knowledge of the future direction of other relay nodes’ movement, based on which they can *cooperate* to make selective and more efficient replication toward the destination. Whether such a scheme can be devised for an ad hoc network with mobile source-destination pairs is still an open research challenge.

IV. CRITICAL DELAY AND 2-HOP DELAY UNDER HYBRID RANDOM WALK MODELS

In this section, we study the critical delay and the 2-hop delay under hybrid random walk models. We first study the critical delay.

A. Critical Delay

Recall that the critical delay is the minimum delay that must be tolerated in order to achieve a throughput of $\omega(1/\sqrt{n})$. Next, we develop lower bounds on the critical delay for the hybrid random walk models. Observe that for $\beta = 0$ (which corresponds to the i.i.d. mobility model) we have the trivial lower bound of 1 on the critical delay. Further, it has been shown in [13] that a throughput of roughly $\Theta(1/n^{1/3}) = \omega(1/\sqrt{n})$ can be achieved with a constant average delay under the i.i.d. mobility model, which implies that 1 is also an asymptotically tight bound (in order terms).

In the sequel, assume therefore that $\beta > 0$. The main idea is to show that if the average delay is below a certain value, then on average the packets travel a distance of $\Theta(1)$ using wireless transmissions before getting delivered to their corresponding destination nodes; and then to show that under the protocol model, this results in a throughput of $O(1/\sqrt{n})$.

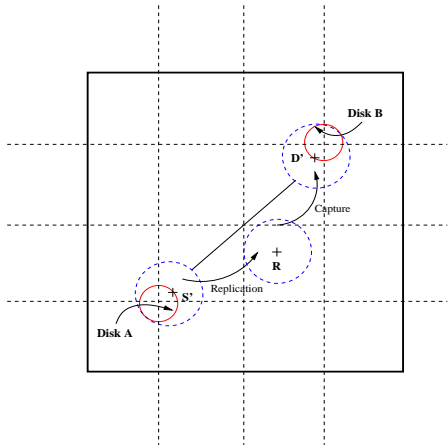


Fig. 4. The figure shows Disks A and B of radius $1/16$ centered at $(-1/4, -1/4)$ and $(1/4, 1/4)$, respectively, that are used in the proof of Lemma 3.

Recall the mobility model described in section II-B. Let $\tau_{E,\beta}^r$ be the first exit time in case of a hybrid random walk model with parameter β . We start by establishing the following lower bound on $\tau_{E,\beta}^{1/8}$, that holds with probability approaching 1 as $n \rightarrow \infty$. The proof is available in Appendix IX-A.

Lemma 2. For $0 < \beta \leq 1/2$, we have

$$\mathbb{P}\left(\tau_{E,\beta}^{1/4} \leq \frac{n^{2\beta}}{1024 \log n}\right) \leq \frac{4}{n^2}.$$

Now we are ready to show that if the average delay is smaller than $\frac{f_o n^{2\beta}}{2048 \log n}$, where $f_o = 1/2400$, then, on average, the packets must be relayed over $\Theta(1)$ distance.

Lemma 3. Consider a network with n nodes moving in accordance with a hybrid random walk model with parameter $\beta > 0$. If the average delay of the packets under a scheduling scheme is smaller than $\frac{f_o n^{2\beta}}{2048 \log n}$, then there exists $N_o < \infty$ such that for all $n \geq N_o$ we have that the packets must, on average, be relayed over a distance no smaller than $f_o/10$.

Proof: Consider disks A and B of radius $1/16$, as shown in Fig.4. Note that according to our assumption in section II-A, the arrival process at each node is independent of the node mobility process. For a given number of nodes n , let f_n denote the fraction of packets having their sources inside disk A and the destinations inside disk B, at the time of arrival. Since the nodes remain uniformly distributed within the subcells at all times under our settings, we have that $f_n \rightarrow (\pi(\frac{1}{16})^2)^2$ as $n \rightarrow \infty$. Thus for all $n \geq N_o$, where N_o is chosen appropriately, we have $f_n \geq f_o$. Now since the average delay is smaller than $\frac{f_o n^{2\beta}}{2048 \log n}$, the delay of at least of one-half of such packets must be at most

$$\frac{2}{f_n} \cdot \frac{f_o n^{2\beta}}{2048 \log n} \leq \frac{n^{2\beta}}{1024 \log n},$$

for all $n \geq N_o$. Consider one such packet; let its time of arrival be t . Using Lemma 2 and the union bound, we have that with

a probability of at least

$$1 - n \cdot \frac{4}{n^2} = 1 - \frac{4}{n},$$

none of the n nodes in the network will exit the disk of radius $1/8$ centered around its initial position at time t before time $t + \frac{n^{2\beta}}{1024 \log n}$. Let S' and D' be the positions of the source and the destination, respectively, for which the distance between the source and the destination at time t is minimized (see Fig.4). Let the destination be captured by the relay node r , and let R be its position when it receives the packet in consideration (see Fig.4). Assuming that the total time spent in the replication (which resulted in the relay r getting a copy of the packet) and capture is $o(n^{2\beta}/\log n)$, the contribution of node movement during the replication and capture phase in carrying the packet toward the destination would be $o(1)$, and can be ignored. Now a simple geometrical argument shows that in order to reach the destination, the packet must be relayed over a distance of at least

$$\frac{1}{\sqrt{2}} - 2 \cdot \frac{1}{16} - 2 \cdot \frac{1}{8} - \frac{1}{8} = d_0 > \frac{1}{5},$$

where the term $\frac{1}{\sqrt{2}} - 2 \cdot \frac{1}{16}$ corresponds to the minimum possible distance between the source and the destination at time t , that is, the distance between the points S' and D' ; the term $2 \cdot \frac{1}{8}$ corresponds to the maximum distance the source and the destination can possibly travel toward each other between time $[t, t + \frac{n^{2\beta}}{1024 \log n}]$; and the term $\frac{1}{8}$ corresponds to the maximum distance that the relay node can possibly move in the direction of the destination after it receives the packet and before time $t + \frac{n^{2\beta}}{1024 \log n}$.

By choosing N_o large enough, we can ensure that $d_0(1 - 1/n) \geq 1/5$ for $n \geq N_o$, and the average distance that the packets must be relayed over in that case would be no smaller than

$$\frac{f_o}{2} \cdot \frac{1}{5} = \frac{f_o}{10},$$

proving the Lemma.

Remark 5. Note that in the above proof we assumed that the total time spent in the replication and capture phase is $o(n^{2\beta}/\log n)$. This assumption is motivated by the fact that in most scheduling schemes the replication and capture are performed using a wireless broadcast or multihop relaying, and are therefore carried out at a much faster time-scale than the node mobility. Note also that a packet might be relayed over several hops during either the replication or capture phase, and for technical consistency, one might then need to scale down the packet size in order to keep the total time spent in replication and capture phases small enough. This kind of packet size scaling has been used quite often in the literature (see, for example, [10], [9], [13], [15]).

We now recall the following result from [14], which shows that if the packets are on average relayed over $\Theta(1)$ distance, then the throughput capacity must be $O(1/\sqrt{n})$.

Lemma 4. Suppose that there exists a constant $c > 0$,

independent of n , such that on average the packets are relayed over a total distance no less than c , then $\lambda(n) = O(1/\sqrt{n})$.

Remark 6. In [14], the above result is proved under the protocol model discussed before in section II-A. By following the line of analysis used for proving Theorem 4.2 in [3], it is possible to establish the above result under the generalized physical model (see [3]), which is more realistic than the protocol model.

The following result is an easy consequence of Lemmas 3 and 4, and the definition of critical delay.

Proposition 1. Under the class of scheduling schemes satisfying Assumption A, the critical delay for a hybrid random walk model with parameter $\beta > 0$ scales as $\Omega(n^{2\beta}/\log n)$.

We now derive an upper bound on the critical delay. Note that the delay under any scheme that can provide a throughput of $\omega(1/\sqrt{n})$ is essentially an upper bound on the critical delay. For $\beta = 1/2$ (which corresponds to the random walk mobility model) it is claimed in [9] that a simple modification of the 2-hop relaying scheme of Grossglauser and Tse can provide a throughput capacity of $\Theta(1)$, incurring a delay of $\Theta(n \log n)$. This result immediately establishes an upper bound of $\Theta(n \log n)$ on the critical delay for $\beta = 1/2$. In what follows, we consider hybrid random walk model with parameter β , for $0 < \beta < 1/2$, and show that the critical delay for the model scales as $O(n^{2\beta} \log n)$.

The idea is to develop a scheduling and relaying scheme that can provide a throughput of $\omega(1/\sqrt{n})$, while incurring a delay of $O(n^{2\beta} \log n)$. In this paper, we only provide the main insight behind such a scheme, leaving out the detailed analysis for our future work.

Consider a scheme in which each packet is replicated at a single relay node, which delivers the packet to the destination node once it is in the same cell as the destination node, possibly using multihop transmission. Note that such a scheme would require an appropriate scaling of the packet size to ensure that the packet can be delivered to the destination within one slot, i.e., before the relay and the destination can possibly move into different cells. We now provide an approximate analysis of the delay and the throughput for such a scheme. In order to keep our discussion simple and insightful, we will ignore possible delays due to queueing at the source or the relay nodes. More precisely, we will assume that the delay of the packet is the time it takes for the relay node to move into the same cell as the destination node, starting from the time it receives the packet from the source node.

Now consider a packet arrival at the source node. Note that since the packet arrival process at each node is independent of the node mobility process, the source node and the destination node are equally likely to be in any of the cells, at the time of the packet arrival. Thus the delay of the packet is of the same order as the expected first hitting time of a single state, in case of a random walk on a 2-d torus of size $n^\beta \times n^\beta$, which, by Lemma 1 is $\Theta(n^{2\beta} \log n)$.

Now for estimating the throughput, we need to account

for the following factors: (i) the loss in throughput due to multiple relayings of the same packet, (ii) the loss in throughput due to the interference. Using the standard multihop scheme, where each hop carries the packet over $\Theta(\sqrt{\log n/n})$ distance[‡], and noting that the packet travels a distance of no more than $\Theta(1/n^\beta)$ using multihop transmissions, it follows that the number of times a packet must be relayed is $O(n^{1/2-\beta}/\sqrt{\log n})$. The loss in throughput due to multiple relayings is correspondingly $O(n^{1/2-\beta}/\sqrt{\log n})$. Now since each transmission is carried over $\Theta(\sqrt{\log n/n})$ distance, it follows easily using the protocol model that the nodes which are within $\Theta(\sqrt{\log n/n})$ distance of the sender must be kept quiet, resulting in a loss of throughput by a factor of $\Theta(\log n)$. Thus the throughput of such a scheme would be $\Omega(n^{\beta-1/2}/\sqrt{\log n}) = \omega(1/\sqrt{n})$ for $\beta > 0$.

The above discussion shows that the critical delay scales as $O(n^{2\beta} \log n)$ for a hybrid random walk model of parameter β . Although, the above arguments are heuristic, they can easily be made precise. However, in order to do so, one would need to specify the details of the scheduling scheme, which is beyond the scope of this paper.

Remark 7. By choosing the size of the capture neighborhood appropriately (e.g., $\Theta(1/\log n)$), one can show that the critical delay is bounded by $\Theta(n^{2\beta} \log \log n)$.

B. 2-Hop Delay

In this section, we analyze the 2-hop delay under the hybrid random walk models. Recall that the original 2-hop relaying protocol of Grossglauser and Tse [4] allows only the nearest neighbor transmissions. Subsequent works [8], [11], [9], [15], [12] have considered a slightly different version of the protocol in which the transmissions are carried out between nodes that are either in the same cell (of size $\Theta(1/\sqrt{n}) \times \Theta(1/\sqrt{n})$), or are within a distance of $\Theta(1/\sqrt{n})$ from each other. Note that these different versions of the protocol are roughly the same, because when n nodes are distributed uniformly within a unit square the nearest neighbor distance is on the order of $\Theta(1/\sqrt{n})$.

Next, we analyze the delay under the 2-hop relaying protocol, assuming that the transmissions are scheduled between nodes that are in the same subcell. As in the analysis of critical delay, we will ignore the queueing delays, postponing their analysis to future work. Thus we would mainly be interested in estimating the time it takes for the relay node and the destination node to come within the same subcell, starting from two randomly and uniformly chosen subcells in the network. Let us denote this random time by T . In subsequent analysis, we will say that two nodes are in a “meeting” if they are currently inside the same cell, and will denote the time between successive meetings as the *inter-meeting time*.

Observe that

$$T = \tau_1 n^{2\beta-1} + \dots + (\tau_1 + \dots + \tau_i) (1 - n^{2\beta-1})^{i-1} n^{2\beta-1} + \dots, \quad (3)$$

[‡]Note that this is the minimum possible order of the communication range for ensuring almost sure connectivity (see [1]).

where τ_1 is the time required for the nodes to enter the same cell, starting from their initial random and uniformly distributed positions, henceforth denoted by *first meeting time*; and τ_i for $i \geq 2$ are the successive *inter-meeting times*. Observe that $n^{2\beta-1}$ is the probability that the nodes choose the same subcell inside a given cell. It is easy to see that the mean first meeting time is of the order of the mean first hitting time of a single state, in case of a random walk on a 2-d torus of size $n^\beta \times n^\beta$. Using Lemma 1, it follows that $\mathbb{E}\{\tau_1\} = \Theta(n^{2\beta} \log n)$. Further, the mean inter-meeting times are of the order of the mean *first return time* (for definition, see, for example, [17, Chap 2, p. 2]) of a random walk on a 2-d torus of size $n^\beta \times n^\beta$, which is well known to be $n^{2\beta}$. We therefore have $\mathbb{E}\{\tau_i\} = \Theta(n^{2\beta})$ for $i \geq 2$. Taking the expectations on both sides of Eq. (3), and performing some simple algebraic manipulations, we obtain

$$\begin{aligned} \mathbb{E}\{T\} &= \mathbb{E}\{\tau_1\} + \mathbb{E}\{\tau_2\}n^{1-2\beta} \\ &= \Theta(n^{2\beta} \log n) + \Theta(n^{2\beta})n^{1-2\beta} \end{aligned}$$

Thus for $\beta < 1/2$, we have $\mathbb{E}\{T\} = \Theta(n)$; and for $\beta = 1/2$, we have $\mathbb{E}\{T\} = \Theta(n \log n)$.

Remark 8. Note that our results for $\beta = 0$ and $\beta = 1/2$ are in agreement with the corresponding results for the i.i.d. mobility model in [8] and the random walk model in [15]. Both these works also account for the queueing delays: In [8], queueing delays at the source nodes as well as relay nodes are considered, whereas, [15] considers queueing delays at the relay nodes only. It is interesting to see that our simplified analysis yields exact results (in order sense) for two extreme choices of β , i.e., $\beta = 0, 1/2$.

Remark 9. Observe that all hybrid random walk models incur roughly $\Theta(n)$ delay under the 2-hop relaying scheme, but their critical delays vary significantly. More precisely, as β increases the critical delay increases as well (roughly as $\Theta(n^{2\beta})$), shrinking the delay-capacity trade-off region. The two extreme cases being: (i) the i.i.d. mobility model (i.e., $\beta = 0$), for which $\omega(1/\sqrt{n})$ capacity can be achieved even under a constant delay constraint; and (ii) the random walk model (i.e., $\beta = 1/2$), for which the delay on the order of $\Theta(n/\log n)$ or more must be tolerated in order to achieve a capacity of $\omega(1/\sqrt{n})$.

V. CRITICAL DELAY AND 2-HOP DELAY UNDER DISCRETIZED RANDOM DIRECTION MODELS

In this section, we study the critical delay and the 2-hop delay under discretized random direction models. As in the previous section, we first study the critical delay.

A. Critical Delay

Recall that the discretized random direction models are characterized by a single parameter α that takes values between 0 and 1/2. As in the previous section, we will first derive a lower bound on the critical delay by lower bounding the first exit time for a disk of radius 1/8. Let $\tau_{E,\alpha}^{1/8}$ denote the first exit time for such a disk in case of discretized random

direction model with parameter α . Let the duration of a time slot be $Cn^{1/2-\alpha}$. For $\alpha = 0$, one trivially obtains a lower bound of $\Theta(\sqrt{n})$ on $\tau_{E,\alpha}^{1/8}$. For $\alpha > 0$, we have the following result, the proof of which follows mutatis mutandis from the proof of Lemma 2:

Lemma 5. For $0 < \alpha \leq 1/2$, we have

$$\mathbb{P}\left(\tau_{E,\alpha}^{1/4} \leq \frac{Cn^{1/2+\alpha}}{1024 \log n}\right) \leq \frac{4}{n^2}.$$

It is interesting to note that a similar result can also be proved for random direction models (see Appendix IX-B):

Lemma 6. Let $T_{E,\alpha}^{1/8}$ denote the first exit time of a disk of radius 1/8 for a random direction model with parameter α . For $0 < \alpha \leq 1/2$, we have

$$\mathbb{P}\left(T_{E,\alpha}^{1/8} \leq \frac{Cn^{1/2+\alpha}}{768 \log n}\right) \leq \frac{4}{n^2}.$$

The following Lemma and Proposition can now be proved in a similar fashion to Lemma 3 and Proposition 1, respectively.

Lemma 7. Consider a network with n nodes moving in accordance with a discretized random direction model with parameter $\alpha > 0$. If the average delay of the packets under a scheduling scheme is smaller than $\frac{Cf_o n^{1/2+\alpha}}{2048 \log n}$, then there exists $N_o < \infty$ such that for all $n \geq N_o$ we have that the packets must on an average be relayed over a distance greater than $f_o/10$.

Proposition 2. Under the class of scheduling schemes satisfying Assumption A, the critical delay for a discretized random direction model with parameter $\alpha > 0$ scales as $\Omega(n^{\alpha+1/2}/\log n)$.

Remark 10. Analogs of the results in Lemma 7 and Proposition 2 can easily be proved for random direction models using Lemma 6.

Remark 11. Note that for $\alpha = 0$, using the lower bound of $\Theta(\sqrt{n})$ on $\tau_{E,\alpha}^{1/8}$, and arguing as in the proof of Lemma 7, we can easily establish a lower bound of $\Theta(\sqrt{n})$ on the critical delay. Moreover, the same reasoning shows that a lower bound of $\Theta(\sqrt{n})$ on critical delay also holds under the random waypoint mobility model. This result was earlier shown in [12], but under a more restricted class of scheduling and relaying schemes than in this paper.

Next, we will establish an upper bound on the critical delay. Consider the scheme discussed before in section IV-A. Recall that the packet is replicated to at most one relay node, which delivers it to the destination node on entering the same cell as the destination node. The approximate analysis of the throughput and delay under such a scheme can be carried out following the line of analysis in section IV-A, and it is straightforward to show that the delay under such a scheme will be $\Theta(n^{1/2+\alpha} \log n)$ and the throughput will be $\Omega(n^{\alpha-1/2}/\sqrt{\log n}) = \omega(1/\sqrt{n})$ for $\alpha > 0$. Thus for $\alpha > 0$, the critical delay is bounded above by $\Theta(n^{1/2+\alpha} \log n)$.

Remark 12. *One might think that by increasing the size of the capture neighborhood to $1/n^\gamma$ where $0 < \gamma < \alpha$, one might be able to reduce the delay below $\Theta(n^{1/2+\alpha} \log n)$, while maintaining a throughput of $\Omega(1/\sqrt{n})$. This is, however, not possible. In fact, it can be shown that with a capture neighborhood of size $r(n)$ the delay becomes $\Theta(n^{1/2+\alpha} \log(1/r(n)))$, and the throughput becomes $\Theta(1/r(n)\sqrt{n \log n})$. Thus choosing a capture neighborhood of size $1/n^\gamma$ for any $\gamma > 0$ will not change the order of the delay. Note also that by increasing the size of the capture neighborhood to $\Theta(1/\log n)$, one can show an upper bound of $\Theta(n^{1/2+\alpha} \log \log n)$ on the critical delay.*

Next, we consider $\alpha = 0$. Note that for $\alpha = 0$, the discretized random direction model is almost the same as the random way-point mobility model, with the only difference being that the successive trips (moving between the chosen pair of points) that a node makes under the discretized random direction model are independent; whereas, there is some dependency between successive trips in case of the random way-point mobility model (since the next trip starts from the point where the previous trip ends). The random way-point mobility model has been analyzed[§] in [12]. In particular, a protocol that allows one to trade-off throughput for the delay has been developed in [12], and shown to achieve the following delay-capacity trade-off:

$$D(n) = O(n/k(n) \log n), \text{ and } \lambda(n) = \Omega(1/k(n) \log n),$$

where $D(n)$ is the average packet delay and the $\lambda(n)$ is the per-node throughput. Following the line of analysis in [12], one can show that the same delay-capacity trade-off can also be achieved under the discretized random direction model with $\alpha = 0$. Now by choosing $k(n) = \sqrt{n}/a(n) \log n$ where

$$a(n) \rightarrow \infty \text{ as } n \rightarrow \infty, \quad (4)$$

it follows that the critical delay is $O(a(n)\sqrt{n} \log^2 n)$ for all $a(n)$ satisfying condition (4). In particular, the critical delay is $o(n^\gamma)$ for any $\gamma > 1/2$.

B. 2-Hop Delay

In this section, we analyze the 2-hop delay under the discretized random direction models. We will assume that the transmissions are scheduled between nodes that are within a distance of $1/\sqrt{n}$ from each other, and will ignore the queueing delays. Thus we would mainly be interested in estimating the time it takes for the relay node and the destination node to come within a distance of $1/\sqrt{n}$ of each other, starting from two randomly and uniformly chosen positions in the network.

Let us denote this random time by T . Arguing as in section IV-B, it can be shown that $\mathbb{E}\{T\} = \Theta(n^{1/2-\alpha})\Theta(\mathbb{E}\{\tau_1\} + \mathbb{E}\{\tau_2\}/p)$, where τ_1 is the first meeting time; τ_2 is the inter-meeting time; and p is the probability that two nodes will come within a distance of $1/\sqrt{n}$ of each other any time during a

[§]Although [12] considers a slightly different version of the random way-point mobility model on a sphere, the results in [12] can easily be extended to a 2-d torus.

time slot, given that they are within the same cell in that time slot. Note that the factor of $\Theta(n^{1/2-\alpha})$ comes because the duration of each time slot is now $\Theta(n^{1/2-\alpha})$. As in section IV-B, we have $\mathbb{E}\{\tau_1\} = \Theta(n^{2\alpha} \log n)$ and $\mathbb{E}\{\tau_2\} = \Theta(n^{2\alpha})$. Furthermore, it is easy to see that $p = \Theta(n^{\alpha-1/2})$. Thus

$$\mathbb{E}\{T\} = \Theta(n^{1/2+\alpha} \log n) + \Theta(n).$$

Hence we see that $\mathbb{E}\{T\} = \Theta(n \log n)$ for $\alpha = 1/2$, and $\Theta(n)$ for $0 \leq \alpha < 1/2$.

Remark 13. *Once again, we note that our results for $\alpha = 0$ and $\alpha = 1/2$ are in agreement with the corresponding results for the random walk model in [15] and the random way-point mobility model in [12]. Both these works also account for the queueing delays: In [12], queueing delays at the source nodes as well as relay nodes are considered, whereas, [15] considers queueing delays at the relay nodes only. Again, we see that our simplified analysis yields exact results (in order sense) for two extreme choices of α , i.e., $\alpha = 0, 1/2$.*

Remark 14. *Observe that all discretized random direction models incur roughly $\Theta(n)$ delay under the 2-hop relaying scheme; however, their critical delays vary significantly. More precisely, as α increases the critical delay increases as well (roughly as $\Theta(n^\alpha)$), shrinking the delay-capacity trade-off region. The two extreme cases being: (i) $\alpha = 0$ (random way-point mobility model), for which $\omega(1/\sqrt{n})$ capacity incurring delays of about $\Theta(\sqrt{n})$; and (ii) $\alpha = 1/2$ (random walk model), for which the delay on the order of $\Theta(n/\log n)$ or more must be tolerated in order to achieve a capacity of $\omega(1/\sqrt{n})$.*

VI. DISCUSSION

The main contribution of this paper is the definition and study of the notion of critical delay, which provides us with a platform to compare and contrast several existing mobility models. The notion of critical delay is important as it provides us a way of determining whether a particular form of node mobility can be exploited to improve the throughput capacity under a given delay constraint. We also showed that there exists a strong connection between the notion of exit time and critical delay, and used this connection to estimate the critical delay under various mobility models.

The results obtained in the previous sections are summarized in Fig.5. Clearly, the mobility models considered in the literature are in some sense extreme: they either exhibit the smallest critical delays or the largest critical delays among all mobility models having roughly the same 2-hop delay. Thus, on one extreme, there is almost no delay-capacity trade-off under the Brownian motion model and the random walk model, and, on the other extreme, there is a smooth delay-capacity trade-off for a wide range of delays under the random way-point mobility model and the i.i.d. mobility model.

An interesting insight provided by our results is that the critical delay is inversely proportional to the *characteristic path length*. By *characteristic path length*, we mean the distance that a node travels without changing direction. (Recall that in

case of (discretized) random direction model with parameter α , the *characteristic path length* is of the order of $n^{-\alpha}$ and the critical delay is roughly of the order of $n^{1/2+\alpha}$.) Thus in terms of application support, a scenario where the nodes move over long distances without changing directions (as in the random way-point mobility model) is more desirable than a scenario where nodes change directions over short distances (as in the Brownian motion model). This is because the former scenario provides more flexibility in terms of choosing the point of operation on the delay-capacity trade-off curve, and can therefore support a wider range of applications.

In a real world scenario, it is rather unlikely that the (density) number of nodes in the network will have a strong influence on the motion of the nodes. We therefore believe that a mobility model like the random way-point model might be more appropriate for determining the scaling laws for large mobile ad hoc networks, rather than a mobility model like the Brownian motion model (random walk model). We therefore expect that future mobile ad hoc networks would provide network designers with ample flexibility in terms of choosing the desired operational point on the delay-capacity trade-off curve, and this opportunity must be fully exploited for optimal operation of such networks, possibly using a cross-layer design approach.

VII. CONCLUSION

We have studied the delay-capacity trade-offs in mobile ad hoc networks. We introduced the meaningful notion of critical delay to systematically study how much delay must be tolerated for a given form of node mobility to result in an improvement of the network capacity. The notion of critical delay allowed us to look at various forms of node mobility studied in the literature from a common perspective, and to compare and contrast them.

We proposed two different classes of mobility models and showed that they both exhibit critical delays that are in-between that of the mobility models studied in the literature, thus showing that the mobility models considered in the literature are rather extreme. More importantly, we showed that the critical delay is inversely proportional to the characteristic path length, which is the distance nodes travel without changing directions. These results, among other things, provide a clear understanding of why is it that the critical delay under the Brownian motion model is larger than the critical delay under the random way-point mobility model.

In a real world scenario, one would expect the number of nodes or the density of nodes to have little, if any, influence on the motion of nodes. Correspondingly, one would expect the characteristic path length to have a rather weak dependence, if at all, on the number of nodes or the node density. One might therefore expect the critical delay in a real world scenario to be close to $\Theta(\sqrt{n})$, as in the case of the random way-point mobility model. This result is optimistic, since it suggests that the future mobile ad hoc networks would provide network designers with ample flexibility in terms of choosing the desired operational point on the delay-capacity trade-off

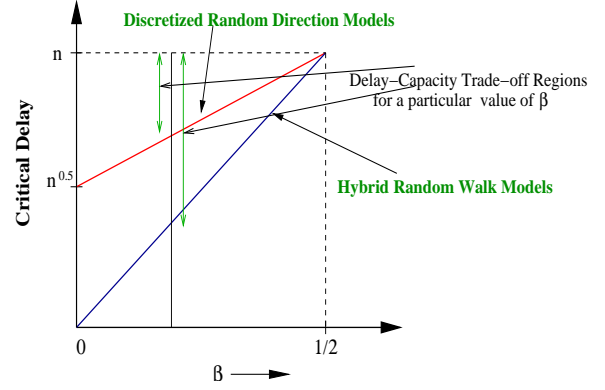


Fig. 5. The figure shows the scaling of critical delay in case of the hybrid random walk models and discretized random direction models.

curve; an opportunity that must be fully exploited for optimal operation of such networks, possibly using a cross-layer design approach.

VIII. ACKNOWLEDGEMENT

We wish to thank Xiaojun Lin for many valuable discussions throughout the course of this work.

IX. APPENDIX

In this section, we provide proofs for Lemmas 2 and 6. We start with the following simple result that is a version of Hoeffding's Inequality (see, for example, [18, Chapter 3, pg. 120]).

Lemma 8. *Let X_1, X_2, \dots, X_n be i.i.d. random variables taking values in $[-l, l]$ for $0 < l < \infty$, and suppose $\mathbb{E}\{X_i\} = 0$ for all i . Let $S_n = \sum_{i=1}^n X_i$ and $\sigma_{S_n}^2$ be the variance of S_n . Then*

$$\mathbb{P}(S_n \geq \mu \sigma_{S_n}) \leq e^{-\mu^2/4},$$

for all $0 \leq \mu \leq 2\sigma_{S_n}/l$.

We are now ready to prove Lemmas 2 and 6.

A. Proof of Lemma 2

Recall the hybrid random walk model of section II-B, and the definition of $\tau_{E,\beta}^{1/8}$, given in section III. As discussed in section III, the statistical properties of the first exit time do not depend on the choice of y . So let y be the origin, that is, the point $(0, 0)$. Let (x_0, y_0) be the cell containing the origin. Also, let (x_t, y_t) be the cell in which node i lies at time t . Further, let

$$\tau_x^+ \triangleq \inf \left\{ t \geq 0 : (x_t - x_0) \geq \frac{n^\beta}{16} \right\};$$

$$\tau_x^- \triangleq \inf \left\{ t \geq 0 : (x_t - x_0) \leq -\frac{n^\beta}{16} \right\};$$

and τ_y^+, τ_y^- be similarly defined with y_t, y_0 in place of x_t and x_0 , respectively. Observe that

$$\mathbb{P}(\tau_{E,\beta}^{1/8} \leq m) \leq \mathbb{P}(\tau_x^+ \leq m \text{ or } \tau_x^- \leq m \text{ or } \tau_y^+ \leq m \text{ or } \tau_y^- \leq m)$$

for $m \geq 0$. Using the union bound and appealing to the symmetry of node motion, we obtain

$$\mathbb{P}(\tau_{E,\beta}^{1/8} < m) \leq 4\mathbb{P}(\tau_x^+ < m).$$

Now observe that before time $\tau_{E,\beta}^{1/8}$, x_t has the following form:

$$x_t = x_0 + \sum_{i=1}^t s_i,$$

where s_i are i.i.d. random variables taking values in $\{-1, 0, 1\}$ with probabilities $\{1/4, 1/2, 1/4\}$, respectively. Although, x_t is not a simple random walk, it is clear due to its symmetry that the reflection principle for 1-d random walk holds in case of x_t as well, and we have

$$\mathbb{P}(\tau_x^+ \leq k) = 2\mathbb{P}\left(x_{\lfloor k \rfloor} - x_0 \geq \frac{n^\beta}{16}\right). \quad (5)$$

for $k \geq 0$, where $\lfloor \cdot \rfloor$ denotes the greatest integer function. Now since each s_i has mean 0 and variance $1/2$, a straightforward application of Lemma 8 gives:

$$\mathbb{P}\left(x_t - x_0 \geq \frac{n^\beta}{16}\right) \leq e^{-\frac{n^{2\beta}}{512t}}, \quad (6)$$

for $t \geq n^\beta/16$. Substituting $k = \frac{n^{2\beta}}{1024 \log n}$ in Eq. (5), and combining with Eq. (6), we obtain

$$\mathbb{P}\left(\tau_x^+ \leq \frac{n^{2\beta}}{1024 \log n}\right) \leq e^{-2 \log n} = \frac{1}{n^2},$$

and the result follows by noting that

$$\mathbb{P}(\tau_{E,\beta}^{1/8} \leq m) \leq 4\mathbb{P}(\tau_x^+ \leq m)$$

for $m \geq 0$.

B. Proof of Lemma 6

Recall the random direction model of section II-B, and the definition of $\tau_{E,\alpha}^{1/8}$, given in section III. Arguing as in the previous proof, it suffices to consider $y = (0, 0)$. Let (x_t, y_t) be the position of the node i after t trips. Let

$$\tau_x \triangleq \inf\{t \geq 0 : |x_t| \geq 1/8\sqrt{2}\},$$

and

$$\tau_y \triangleq \inf\{t \geq 0 : |y_t| \geq 1/8\sqrt{2}\}.$$

It is then clear that

$$\mathbb{P}(\tau_{E,\alpha}^{1/8} \leq m) \leq \mathbb{P}(\tau_x \leq m \text{ or } \tau_y \leq m).$$

Appealing to the symmetry of the node motion and using the union bound, we obtain

$$\mathbb{P}(\tau_{E,\alpha}^{1/8} \leq m) \leq 2\mathbb{P}(\tau_x \leq m).$$

Let s_k be the x-coordinate of the nodes' position immediately after completing the k^{th} trip. Before time $\tau_{E,\alpha}^{1/8}$, s_k has the simple form:

$$s_k = \sum_{i=1}^k z_i,$$

where z_i are i.i.d. random variables taking values in $[-n^{-\alpha}, n^{-\alpha}]$. Note also that each z_i has mean zero and variance $n^{-2\alpha}/2$. Using Lemma 8, we have

$$\mathbb{P}(s_k \geq 1/8\sqrt{2}) \leq e^{-\frac{n^{2\alpha}}{256k}},$$

for $k \geq n^\alpha/8\sqrt{2}$. Using the symmetry of the node motion once again, we have

$$\mathbb{P}(|s_k| \geq 1/8\sqrt{2}) \leq 2e^{-\frac{n^{2\alpha}}{256k}}.$$

Noting that the duration of each trip is $Cn^{1/2-\alpha}$, it follows that

$$\begin{aligned} \mathbb{P}\left(\tau_x \leq k \cdot Cn^{1/2-\alpha}\right) &= \mathbb{P}\left(\bigcup_{i=1}^k |s_i| \geq 1/8\sqrt{2}\right) \\ &\leq \sum_{i=\lceil n^\alpha/8\sqrt{2} \rceil}^k 2e^{-\frac{n^{2\alpha}}{256i}} \\ &\leq k \cdot 2e^{-\frac{n^{2\alpha}}{256k}}, \end{aligned}$$

where $\lceil n^{2\alpha}/8\sqrt{2} \rceil$ denotes the smallest integer greater than $n^{2\alpha}/8\sqrt{2}$. Now since $n^{2\alpha} \leq n$ for $\alpha \leq 1/2$, we have

$$\mathbb{P}\left(\tau_x \leq \frac{Cn^{2\alpha}}{768 \log n} \cdot n^{1/2-\alpha}\right) \leq n \cdot 2e^{-3 \log n} = \frac{2}{n^2}.$$

Thus

$$\mathbb{P}\left(\tau_{E,\alpha}^{1/8} \leq \frac{Cn^{\alpha+1/2}}{768 \log n}\right) \leq \frac{4}{n^2},$$

as claimed.

REFERENCES

- [1] P. Gupta and P.R. Kumar, "The capacity of wireless networks," *IEEE Trans. on Information Theory*, vol. IT-46, no. 2, pp. 388–404, March 2000.
- [2] 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://fleece.ucsd.edu/~massimo/>, 2005.
- [3] 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.
- [4] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad-hoc wireless networks," in *IEEE INFOCOM*, 2001, pp. 1360–1369.
- [5] N. Bansal and Z. Liu, "Capacity, delay and mobility in wireless ad-hoc networks," in *IEEE INFOCOM*, April 2003.
- [6] E. Peravalov and R. Blum, "Delay limited capacity of ad hoc networks: Asymptotically optimal transmission and relaying strategy," in *IEEE INFOCOM*, April 2003.
- [7] A. Tsirigos and Z.J. Haas, "Multipath routing in presence of frequent topological changes," *IEEE Communications Magazine*, Nov. 2001.
- [8] M.J. Neely and E. Modiano, "Capacity and delay tradeoffs for ad-hoc mobile networks," to appear in *IEEE Trans. on Information Theory*, 2005.
- [9] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-delay trade-off in wireless networks," in *IEEE INFOCOM*, 2004.
- [10] S. Toumpis and A. Goldsmith, "Large wireless networks under fading, mobility, and delay constraints," in *IEEE INFOCOM*, 2004.
- [11] G. Sharma and R. Mazumdar, "Scaling laws for capacity and delay in wireless networks with random mobility," in *IEEE ICC*, 2004.
- [12] G. Sharma and R. Mazumdar, "Delay and capacity trade-off in wireless ad hoc networks with random way-point mobility," Preprint, Dept. of ECE, Purdue University, available at <http://ece.purdue.edu/~gsharma/>, 2005.

- [13] 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.
- [14] X. Lin, G. Sharma, R. Mazumdar, , and N. Shroff, "Degenerate delay-capacity trade-offs in ad hoc networks with brownian mobility," Preprint, School of ECE, Purdue University. Available at: <http://ece.purdue.edu/~gsharma>, 2005.
- [15] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-delay trade-off in wireless networks – part i: The fluid model," Preprint, Dept. of EE, Stanford University. Available at: <http://www.stanford.edu/~jmammen/>, 2005.
- [16] P. Nain, D. Towsley, B. Liu, and Z. Liu, "Properties of random direction models," in *IEEE INFOCOM*, March 2005.
- [17] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, Monograph in preparation, available at <http://stat-www.berkeley.edu/users/aldous/RWG/book.html>, 2002.
- [18] A. Gut, *Probability: A Graduate Course*, Springer, 2005.