Title
Mobility-Capacity-delay Trade-off in Wireless Ad Hoc Networks

Journal
Ad Hoc Networks, 4(5)

Author
Garcia-Luna-Aceves, J.J.

Publication Date
2006-09-01

Peer reviewed

ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Ad Hoc Networks xxx (2005) xxx–xxx

Ad Hoc Networks

www.elsevier.com/locate/adhoc

Mobility–capacity–delay trade-off in wireless ad hoc networks

Renato M. de Moraes [a,*], Hamid R. Sadjadpour [a], J.J. Garcia-Luna-Aceves [b]

[a] Department of Electrical Engineering, University of California at Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, United States
[b] Department of Computer Engineering at UCSC and Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304, United States

Received 23 August 2004; received in revised form 18 May 2005; accepted 16 June 2005

Abstract

We show that there is a trade-off among mobility, capacity, and delay in ad hoc networks. More specifically, we consider two schemes for node mobility in ad hoc networks. We divide the entire network by cells whose sizes can vary with the total number of nodes *n*, or whose size is independent of the number of nodes. We restrict the movement of nodes within these cells, calculate throughput and delay for randomly chosen pairs of source–destination nodes, and show that mobility is an entity that can be exchanged with capacity and delay. We also investigate the effect of directional antennas in a static network in which packet relaying is done through the closest neighbor and verify that this approach attains better throughput than static networks employing omnidirectional antennas.

Keywords: Ad hoc networks; Delay; Information theory; Mobility; Throughput capacity

1. Introduction

Capacity analysis in ad hoc networks has become an important issue since Gupta and Kumar [1] showed that the capacity of a fixed and connected wireless network decreases as the number of nodes *n* increases. Grossglauser and Tse [2] presented a two-phase packet forwarding technique for mobile ad hoc networks (MANETs) in which a source node transmits a packet to the nearest neighbor, and that relay delivers the packet to the destination when this destination becomes the closest neighbor of the relay. The scheme was shown [2] to attain constant per source–destination throughput as the number of nodes in the MANET increases by taking advantage that communication among nearest nodes cope the interference due to far nodes. To date, several schemes

* Corresponding author. Tel.: +1 831 459 4141; fax: +1 831 459 4829.

E-mail addresses: renato@soe.ucsc.edu (R.M. de Moraes), hamid@soe.ucsc.edu (H.R. Sadjadpour), jj@soe.ucsc.edu (J.J. Garcia-Luna-Aceves).

1570-8705/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.adhoc.2005.06.005

have traded off delay in order to attain higher capacity in mobile ad hoc networks (MANETs) [2–7].

In this paper, we present new network models to show that mobility can also be traded as a resource together with capacity and delay. The idea is to allow the nodes execute *restricted* movements, i.e., each node moves only inside some given area in the network. By allowing transmissions to closest neighbor nodes only, we overcome interference from other transmitting nodes. Given that nodes have restrained mobility, the delivery from source to destination is done across multiple hops obtained by relaying packets along the path linking the source to the destination. Diggavi et al. [3] considered a restrained one-dimensional mobility model in which nodes were allowed to execute movements on circles on a sphere. They showed that a constant throughput is still feasible; however, they do not present the corresponding trade-off associated to mobility, capacity and delay.

Note that restrained mobility patterns have potential practical applications in those cases that nodes are not allowed to leave a given region like a room, a hallway, or other city predefined in the area the region covered by a sensor network, and has to rely on multiple hops (i.e., relays) to send a packet to a farther destination. Therefore, restricted mobility models are important to the study of ad hoc networks.

As defined by Gupta and Kumar [1], Grossglauser and Tse [2], and Gamal et al. [7], a *node throughput* (or simply throughput) of $\Lambda(n)$ bits/s is feasible if every node can send information at a rate of $\Lambda(n)$ bits/s to its chosen destination. Furthermore, the *delay* $D(n)$ of a packet in a network is the time it takes the packet to reach the destination after it leaves the source, where queuing delay at the source is not considered. The average packet delay for a network with n nodes is obtained by averaging over all packets, all source–destination pairs, and all random network configurations.

Section 2 summarizes the network model that has been used recently to analyze the capacity of wireless network [1,2,5,7]. Section 3 presents a restricted mobility model, which we call *Scheme 1*, where the size of the cells varies with the number of nodes n. The associated throughput ($\Lambda_1(n)$) and delay ($D_1(n)$) as functions of n are given by[1]

$$\Lambda_1(n) = \Theta\left(\sqrt{\frac{\log(n)}{n}}\right) \quad \text{and} \quad D_1(n) = \Theta(\sqrt{n}).$$

Compared to the static network model [1], *Scheme 1* attains a gain of $\Theta(\log(n))$ by using restrained mobility. Section 4 presents another restricted mobility model, which we call *Scheme 2*, in which the size of a cell is not a function of n. For a given constant number of cells l, the size of each cell is $1/l$, and the corresponding throughput ($\Lambda_2(n)$) and delay ($D_2(n)$) are

$$\Lambda_2(n) = \frac{1}{\sqrt{l}}\Theta(1) \quad \text{and} \quad D_2(n) = \Theta\left(\frac{n}{l}\right).$$

This throughput result is a generalization of the results by Grossglauser and Tse [2] and represents a reduction of $1/\sqrt{l}$, while the delivery delay is decreased. This indicates that mobility, capacity, and delay should be treated as exchangeable entities. Section 5 presents a modification of *Scheme 2* to allow multiple-copy relaying [8,9] so that the order of magnitude of the throughput is preserved, but lower delivery delay is attained.

Section 6 presents the throughput-delay analysis for a fixed network in which nodes are endowed with directional antennas. Nodes relay packets to their closest neighbors along the path to destinations. We find that the throughput ($\Lambda_D(n)$) and delay ($D_D(n)$) for this scheme are

$$\Lambda_D(n) = \Theta\left(\sqrt{\frac{\log(n)}{n}}\right) \text{ and } D_D(n) = \Theta\left(\sqrt{\frac{n}{\log(n)}}\right).$$

This result is important, because it represents a capacity gain of $\Theta(\log(n))$ compared to the results in Gupta and Kumar [1] and Yi et al. [10].

[1] Here we use the Knuth's notation: $f(n) = \Theta(g(n))$ means there are positive constants c_1, c_2, and N, such that $0 \leqslant c_1 g(n) \leqslant f(n) \leqslant c_2 g(n)$ $\forall n \geqslant N$. Also, $\log(\cdot)$ stands for the natural logarithm.

2. Basic network model

The model considered here is that of a wireless ad hoc network with nodes assumed either fixed or mobile. The network consists of a normalized unit area torus containing n nodes [1,2,7].

For the case of *fixed nodes*, the position of node i is given by X_i. A node i is capable of transmitting at a given transmission rate of W bits/s to j if [1]

$$|X_k - X_j| \geqslant (1+\Delta)|X_i - X_j|, \quad (1)$$

where $\Delta > 0$, so that node X_k will not impede X_i and X_j communication. This is called the *protocol model* [1].

For the case of *mobile nodes*, the position of node i at any time is now a function of time. A successful transmission between nodes i and j is governed again by Eq. (1), where the positions of the nodes are time dependent [2]. Time is slotted to simplify the analysis. Also, at each time step, a scheduler decides which nodes are sources, relays, or destinations, in such a manner that the association pair, source–destination, does not change with time. Nodes are assumed to move according to a *uniform mobility model* [5]. In this model, the nodes are initially uniformly distributed, and move at a constant speed $v(n)$ and the directions of motion are independent and identically distributed (iid) with uniform distribution in the range $[0, 2\pi)$. As time passes, each node chooses a direction uniformly from $[0, 2\pi)$ and moves in that direction, at a speed $v(n)$, for a distance z, where z is an exponential random variable with mean μ. After reaching z the process repeats. This model satisfies the following properties [5]: (a) the position of the nodes are independent of each other, at any time t, (b) the steady-state distribution of the mobile nodes is uniform, and (c) the direction of the node movement is uniformly distributed in $[0, 2\pi)$, conditional on the position of the node.

3. Scheme 1

We present a restricted mobility scheme that attains a capacity gain of $\Theta(\log(n))$ compared to the static network model [1]. The throughput still

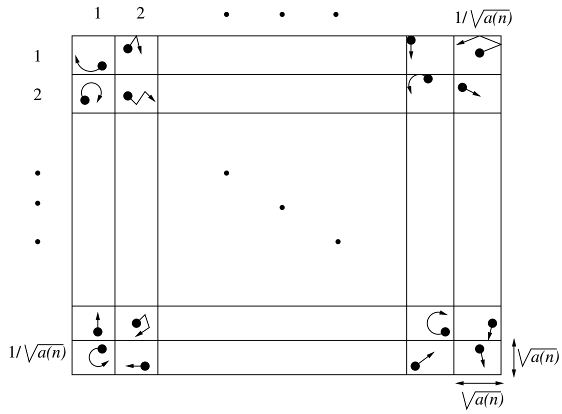

Fig. 1. Unit area torus network divided into $1/a(n)$ cells, each with size of $a(n)$.

decreases as the number of nodes n in the network grows to infinity. However, it serves as a building block for the scheme presented in the next section, which attains non-zero asymptotic throughput capacity in a dense network.

The model we propose is illustrated in Fig. 1. The network is a unit torus divided into square cells, each of area $a(n)$ as in [7], in which they showed that, if $a(n) \geqslant \frac{2\log(n)}{n}$, then each cell has at least one node *with high probability (whp)*, i.e., with probability $\geqslant 1 - 1/n$. This condition guarantees connectivity[2] *whp* [1,7].

We now consider the additional assumption that each node has its movement confined to only one cell. This means that a node cannot cross the cell edge and percolate to a neighbor cell. By doing so, each cell is composed by at least one node *whp*, and such a node moves with speed $v(n)$, and no preferential direction of movement within the cell. Nodes move independently of each other, and once they hit the cell boundaries they are bounced back (with relation to the edge normal).

We assume that each node only communicates with another node from an adjacent cell, and this happens only when the nodes are close enough to each other (i.e., both are near to the common

[2] In [11], the connectivity criterion is relaxed, where percolation theory ensures that a connected backbone transports the total amount of traffic of the network and provides a node throughput of $\Theta(1/\sqrt{n})$.

edge that separates the cells) so that the effect of interference can be minimized. Thus, a source node relies on relays across several cells to have a packet delivered to a destination. Each packet travels via multiple relays from source to destination following the path close to the straight line linking source and destination. Each source–destination pair is chosen uniformly and independently from different cells. Fig. 2 illustrates a packet whose source and destination nodes are in cells i and d respectively, separated by an average distance $\overline{L}$. Possible cell paths for this packet are $\{i\to j\to f\to g\to c\to d\}$, $\{i\to j\to f\to g\to h\to d\}$, $\{i\to e\to f\to g\to c\to d\}$, $\{i\to e\to f\to g\to h\to d\}$, for example.

Grossglauser and Tse [2] showed that transmission to the nearest node is possible, even when the number of interferers in the network scale to infinity. This allows a node to schedule transmission to a neighbor node from an adjacent cell when Eq. (1) is satisfied. In addition, we assume that both nodes are so close that communication is successful during the entire time slot (or session). The transmission is half-duplex so that each node uses half of the communication time slot to transmit at a rate of W bits/s, and the other half to receive at the same rate. Thus, the average available bit rate is $W/2$ bits/s. Each time two nodes communicate with each other, they exchange packets, and these exchanges can be source–relay, relay–relay, or relay–destination transmissions.

The area in which successful communication can occur is shown in Fig. 2. Basically, it is a semi-circumference $b(n)$ of radius $\frac{\sqrt{a(n)}}{2+2\sqrt{2}}$ where two nodes from adjacent cells can come close to each other so that Eq. (1) is satisfied, i.e., no other node from the other cells will be closer to them than themselves. For the case in which more than one node in the same cell are simultaneously traveling inside $b(n)$, only one of these nodes is allowed to communicate with a node from the adjacent cell. Accordingly, from Fig. 2, the two adjacent nodes in cells i and j are able to communicate during the time they simultaneously travel inside their respective regions $b(n)$'s in their cells as shown. We have that

$$b(n)=\frac{1}{2}\pi\left(\frac{\sqrt{a(n)}}{2+2\sqrt{2}}\right)^2=\frac{\pi a(n)}{24+16\sqrt{2}}. \quad (2)$$

The probability of finding a node traveling inside $b(n)$ is $b(n)/a(n)$, because the node has no preferential direction of movement in the cell and tends to move uniformly inside the cell. In addition, because

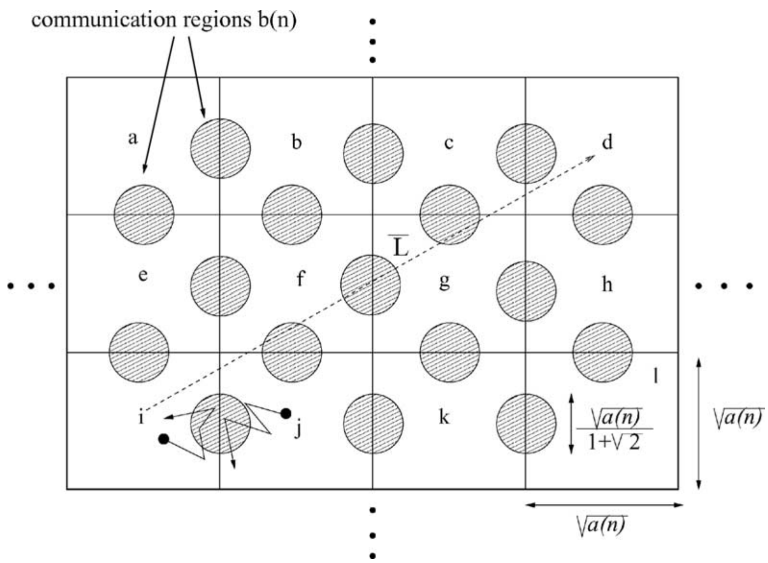

Fig. 2. Region $b(n)$ where communication between nodes from adjacent cells is possible.

the nodes have iid movements, the probability that both nodes come to the communication region simultaneously, denoted by P_c, equals

$$P_c = \left[\frac{b(n)}{a(n)}\right]^2 = \left(\frac{\pi}{24+16\sqrt{2}}\right)^2 = c_3. \quad (3)$$

Hence, P_c does not depend on n.

Because $\overline{L}$ is the mean distance between two uniformly and independently chosen source–destination nodes in the network, the average path distance across cells traversed by a packet from source to destination is $\Theta(\overline{L})$. Accordingly, each cell hop has an average size of $\sqrt{a(n)}$. Thus, the mean number of hops traversed by a packet is $\frac{\Theta(\overline{L})}{\sqrt{a(n)}}$.

According to the definition of throughput, each source generates $\Lambda(n)$ bits/s and there are n sources in the network. Also, each bit needs to be relayed by $\frac{\Theta(\overline{L})}{\sqrt{a(n)}}$ nodes on the average. Thus, the total average number of bits per second served by the entire network equals $\frac{\Theta(\overline{L})n\Lambda(n)}{\sqrt{a(n)}}$. To ensure that all required traffic is carried, we need that

$$c_4 n\frac{W}{2}P_c \leqslant \frac{\Theta(\overline{L})n\Lambda(n)}{\sqrt{a(n)}} \leqslant c_5 n\frac{W}{2}P_c$$
$$\Rightarrow c_6\sqrt{a(n)} \leqslant \Lambda(n) \leqslant c_7\sqrt{a(n)}. \quad (4)$$

We just proved the following theorem.

Theorem 1. *For Scheme* 1 *with* $a(n) = \frac{k\log(n)}{n}$ *and* $k \geqslant 2$, *to guarantee connectivity*, *we have*

$$\Lambda_1(n) = \Theta\left(\sqrt{\frac{\log(n)}{n}}\right).$$

Compared to the capacity result obtained by Gupta and Kumar [1] which is $\Theta(1/\sqrt{n\log(n)})$, the result of Theorem 1 represents a gain of $\Theta(\log(n))$. *Thus, by allowing the nodes to execute a restricted mobility pattern we obtain a throughput gain over the static network model.*

Although in this model we have used mobility and multi-user diversity [12] to overcome interference (note that Gupta and Kumar [1] could not use multi-user diversity because they consider only fixed nodes), the network still does not scale well with the number of nodes, i.e., $\Lambda_1(n) \to 0$ when n goes to infinity. This happens because the number of hops necessary to reach a destination increases with n, so that the same packet is retransmitted infinite times as n grows to infinity, thus wasting the available bandwidth. The model we present in the next section does not have this problem, and it is indeed a generalization of the results obtained by Grossglauser and Tse [2].

The average delay incurred by a packet to reach the destination in *Scheme 1* is the sum of the average time a packet spends in each hopping cell in the path to its destination. A node travels around the cell boundary on average every $t(n)$ time slots that is proportional to

$$t(n) \propto \frac{\Delta S \cdot P_c}{v(n)} \Rightarrow t(n) = \Theta\left(\frac{\sqrt{a(n)}}{v(n)}\right), \quad (5)$$

where $\Delta S = \Theta(\sqrt{a(n)})$ is the average distance in one-round trip inside a cell. Note also that the total number of hops is $\Theta(\overline{L}/\sqrt{a(n)})$, and that the speed of each node must be a function of n, because we assume that the total network area is constant. To model a real network in which a node would occupy a constant area, if the network grows, the entire area must grow accordingly. Therefore, because in our analysis we maintain the total area fixed, we must scale down the speed of the nodes [7]. Consequently, the velocity of the nodes ($v(n)$) must decrease with $1/\sqrt{n}$. Combining all this information, the average delay (D_1) in *Scheme 1* is

$$D_1(n) = (\#\text{ of hops}) \cdot t(n) = \Theta\left(\frac{1}{v(n)}\right) = \Theta(\sqrt{n}). \quad (6)$$

This delay is larger than that obtained by Gupta and Kumar [1], which was shown to be $\Theta(1/\sqrt{a(n)}) = \Theta(\sqrt{n/\log(n)})$ [7]. This is a direct consequence of the throughput-delay trade-off property [7]. *The capacity improvement is obtained at the cost of increase in delay.*

4. Scheme 2

In the previous section we saw that, by having an infinite number of relays (or hops), the capacity of the network decreases as the number of nodes increases. Here, we show that, by having a finite number of relays and using local transmission to overcome interference, we can attain constant throughput as n increases, but we can also trade-off the number of hops with capacity and delay, i.e., we can exchange mobility by capacity and delay, which is a generalization of the results by Grossglauser and Tse [2].

Fig. 3 shows the network and its cells. Now, the network area is divided into l square cells and l is a network design parameter that does not depend on n. Hence, each cell has area of size $1/l$. Again, we assume that the n nodes are uniformly distributed over the entire network, but each node is restricted to move only inside of its cell (one of the l cells). Among the total number of nodes n, a fraction of them, n_S, are randomly chosen as senders, while the remaining nodes, n_R, function like possible receiving nodes [2]. A sender density parameter θ is defined as $n_S = \theta n$, where $\theta \in (0,1)$, and $n_R = (1-\theta)n$. Each node can be a source for one session and a destination for another session. Nodes travel with velocity $v(n)$, have no preferential direction of movements, move independently of each other, and once they hit cell boundaries they bounce back with relation to the edge normal. Here, we consider that each node can communicate with its closest neighbor within the transmission range r_0, whether this neighbor is inside its own cell or from an adjacent cell (when it is traveling around the cell boundary). For a uniform distribution of the nodes, $r_0 = 1/\sqrt{\theta\pi n}$ [8,9]. Thus, communication takes place every time nodes come close enough so that transmission is successful. Moreover, communication between two nodes from the same cell can only be a source–destination, or a relay–destination packet exchange. A relay–relay communication only happens between nodes from different neighboring cells.

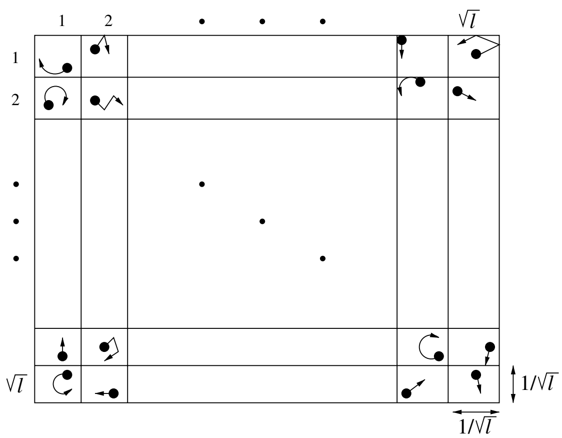

Fig. 3. Unit area torus network divided into l cells, each with size area of $1/l$.

A source–destination pair is uniformly chosen among the n nodes, so that the destination does not have to be necessarily in the same cell as its source. Thus, again, a packet may traverse relays to reach its destination. We assume that, once a packet is relayed to a cell, it is not relayed again for another node in the same cell. Instead, the node keeps the packet in its queue, until it reaches the neighborhood of an adjacent cell in the path toward the destination, so that it forwards the packet to the closest receiver node in the neighboring cell. In this model there is no fixed communication region as in the previous model. Once the node moves close enough around the cell boundary and there is a neighbor receiver node from the adjacent cell moving within the transmission range r_0, then it relays the packet to this neighbor if there is a packet to forward in that direction, so that it can be either a source–relay, or relay–relay, or relay–destination transmission. The communication is simplex, so that each sender node uses the entire communication time slot to transmit at rate W bits/s.

Furthermore, because the nodes move independently of each other, once the network is in steady state, each node in a cell will come closer to another node in that cell at some point in time so that they can exchange packets. This same idea applies to neighbor nodes: because nodes move independently of one another, two nodes from adjacent cells will come close to each other, around the boundary which separates their cells, at some point in time, such that they can exchange packets. Therefore, in steady state, the traffic of each node will be uniformly distributed among neighbors in the same cell, as well as among neighbors from each adjacent cell. Accordingly, for the network in steady state, we have

- Each node has a packet for another node in the same cell.
- Each node has a packet for another node in each of its neighbor cells whose communication is possible.

In addition, for a finite l and a sufficiently large n, connectivity is guaranteed if $\frac{1}{l} > \frac{2\log(n)}{n}$ (i.e., the cell size is greater than $2\log(n)/n$), and because of the uniform distribution of the nodes, each cell contains $\Theta(\frac{n}{l})$ nodes. Because $n \to \infty$, l can be chosen to be any positive integer to satisfy connectivity criterion.

As before, $\overline{L}$ is the mean distance between a source and destination uniformly and independently chosen in the network, thus the average path length across cells followed by a packet is $\Theta(\overline{L})$. Given that each cell hop has an average size of $1/\sqrt{l}$, the average number of hops traversed by a packet until it reaches its destination is $\frac{\Theta(\overline{L})}{1/\sqrt{l}}$.

According to the definition of throughput, each source generates $\Lambda(n)$ bits/s, with n_S being sources in the network. Because each bit needs to be relayed on the average by $\frac{\Theta(\overline{L})}{1/\sqrt{l}}$ nodes, the total average number of bits per second served by the entire network equals $\frac{\Theta(\overline{L})n_S\Lambda(n)}{1/\sqrt{l}}$. Hence, to ensure that all required traffic is carried, we need that

$$c_8 n_S W \leqslant \frac{\Theta(\overline{L})n_S\Lambda(n)}{1/\sqrt{l}} \leqslant c_9 n_S W \Rightarrow \frac{c_{10}}{\sqrt{l}} \leqslant \Lambda(n) \leqslant \frac{c_{11}}{\sqrt{l}}. \tag{7}$$

This proves the following theorem.

Theorem 2. *For Scheme* 2, *for finite l and sufficiently large n, we have*

$$\Lambda_2(n) = \frac{1}{\sqrt{l}}\Theta(1).$$

Theorem 2 is a generalization of the results by Grossglauser and Tse [2], given that we have divided the network into l equal cells. If we set $l = 1$, Theorem III.5 in [2] follows.

Because no node is allowed to move through the entire network, a packet stored in the relay queue of a node has to follow a path of cells in the direction of the destination. Therefore, we should expect a smaller delay than that obtained in the scheme by Grossglauser and Tse [2]. The average delay (D_2) in *Scheme 2* is given by the time the packet spends hopping until it reaches the destination cell, plus the amount of time the last relay in the destination cell needs to reach the destination node. The later is $\Theta\left(\frac{n}{l}\right)$, because we have $\Theta\left(\frac{n}{l}\right)$ nodes in each cell [7,6]. The former is given by the number of hops traversed multiplied by the average time spent per hop (i.e., (# of hops) $\cdot\, t(n)$), which is $\Theta\left[\frac{\overline{L}}{1/\sqrt{l}}\left(\frac{1}{v(n)}\frac{1}{\sqrt{l}}\right)\right] = \Theta(\sqrt{n})$. Thus,

$$\begin{aligned} D_2(n) &= \text{delay during hopping} \\ &\quad + \text{delay in destination cell} \\ &= \Theta\left(\sqrt{n} + \frac{n}{l}\right) \approx \Theta\left(\frac{n}{l}\right) \quad (\text{for } n \text{ large}), \end{aligned} \tag{8}$$

because the term n/l dominates $\sqrt{n}$, for a sufficiently large value of n (and $l \ll \sqrt{n}$). Comparing $D_2(n)$ to the delay attained in the scheme by Grossglauser and Tse [2], whose delay was shown to be $\Theta(n)$ [6,7], we conclude that, as we expected, the delay in *Scheme 2* is smaller by a factor of l.

From Theorem 2, Eq. (8), and comparing with [2], we conclude that we can trade-off mobility as a resource with capacity and delay. By restraining the nodes to move inside cells of size area $1/l$, the $\Theta(1)$ throughput obtained in [2] is reduced by a factor of $\sqrt{l}$, while the delivery delay is decreased by a factor of l. Thus, *Scheme 2* is a generalization of the network model by Grossglauser and Tse [2].

The next section presents a modified version of *Scheme 2* that allows more than one copy of a packet to be forwarded at the destination cell, such that lower delivery delay is possible.

5. Scheme 2 with multi-copy relaying at destination cell

We now introduce an improved packet forwarding strategy [8,9] for mobile ad hoc networks that attains the $\Theta(1)$ capacity of the basic scheme by Grossglauser and Tse [2], but provides lower delay.

We maintain all assumptions from *Scheme 2*, but change the last relaying phase in which a node (a sender or relay) from an adjacent cell has to

ARTICLE IN PRESS

forward a packet to the destination cell. Hence, once a relay node reaches the boundary of the destination cell, it forwards at once copies of the packet to multiple one-time relay nodes located at the destination cell that are within the transmission range r_0 of him. By doing so, the time within which a copy of the packet reaches its destination can be decreased in that cell. The first one-time relay node that reaches the destination close enough delivers the packet.

In *Scheme 2*, a relay approaching the destination cell transmits to its nearest receiver neighbor in the destination cell, so that the interference caused by other nodes is low, allowing reliable communication. However, it may be the case that the relay can have more than one receiver neighbor node from the destination cell in the transmission range, and we can take advantage of that. We allow those additional receiving neighbor nodes to also have a copy of the packet. Hence, instead of only one copy, K-copies will follow different random routes in the destination cell and can find the destination node earlier compared to *Scheme 2*. In addition, packets are assumed to have header information for scheduling and identification purposes, and a time-to-live (TTL) threshold field as well. We assume that, before any packet is transmitted between nodes, a handshake takes place at the beginning of the time slot, such that no relay transmits a packet that a destination has already received. In this way we enforce only one-copy delivery. Also, after the TTL expires, the packet is dropped from the additional relaying nodes queues which did not deliver the copy of the packet.

For the sake of completeness, we reproduce below the main results from [8,9], which provides a complete description of the multi-copy technique and its analysis.

5.1. Single-copy forwarding case

Because we have node trajectories independent and identically distributed, we focus on a given relay node labeled as *node 1* at the destination cell, and without loss of generality assume that *node 1* received a packet from a relay moving in the boundary of the neighbor cell during time $t_0 = 0$. Let $P\{|X_1(s) - X_{\text{dest}}(s)| \leqslant r_0|s\}$ denote the probability that relay *node 1* at position $X_1(s)$ is close enough to the destination node dest given that the time interval length is s, where r_0 is the relay transmission radius so that successful delivery is possible. The time interval length s is the delivery-delay random variable accounted in the destination cell. Perevalov and Blum [4] obtained an approximation for the ensemble average with respect to all possible uniformly-distributed starting points, $(X_1(0), X_{\text{dest}}(0))$, where they considered the nodes moving on a sphere. We can extend their result to nodes moving on the torus and have [4]

$$
\begin{aligned}
E_{\text{U}}[P\{|X_1(s) - X_{\text{dest}}(s)| \leqslant r_0|s\}] &= 1 - \mathrm{e}^{-\lambda s} \\
&\quad \cdot \left(1 - \lambda \mathrm{e}^{-\lambda\int_0^s h_{X'}(t)\,\mathrm{d}t} \int_0^s \mathrm{e}^{\lambda\int_0^t h_{X'}(u)\,\mathrm{d}u} h_{X'}(t)\,\mathrm{d}t\right) \\
&= P\{S \leqslant s\} = F_S(s),
\end{aligned} \tag{9}
$$

where $E_{\text{U}}[\cdot]$ means the ensemble average over all possible starting points that are uniformly distributed on the torus. $F_S(s)$ can be interpreted as the cumulative density function of the delay random variable S. The function $h_X(t)$ is the difference from the uniform distribution, such that $h_X(0) =$!0 and $|h_X(t)| < 1$ for all t, and X' is a point at distance r_0 from the destination. The parameter λ is related to the mobility of the nodes and can be expressed by [4]

$$
\lambda = \frac{2r_0v}{\pi R^2} = \frac{2r_0v}{1} = 2r_0v, \tag{10}
$$

which results from evaluating the flux of nodes entering a circle of radius r_0 during a differential time interval considering the nodes uniformly distributed over the entire torus of unit area and traveling at speed $v(n)$. For a uniform distribution of the nodes $r_0 = 1/\sqrt{\theta\pi n}$. Hence, the radius r_0 decreases with $1/\sqrt{n}$. As before, the velocity of the nodes decreases with $1/\sqrt{n}$. Then

$$
\lambda = \frac{1}{\Theta(n)}. \tag{11}
$$

Now, $h_X(t)$ has to be taken according to the random motion of the nodes [4]. If we consider the *uniform mobility model* [5], then a steady-state uniform distribution results as the random motion of the nodes on the torus. In such a case, $h_X(t) = 0$ $\forall t \geqslant 0$. Applying this result in Eq. (9) we have

$$E_U[P\{|X_1(s) - X_{\text{dest}}(s)| \leqslant r_0|s\}] = 1 - e^{-\lambda s} = P\{S \leqslant s\} = F_S(s), \tag{12}$$

which has the following probability density function:

$$f_S(s) = \frac{dF_S}{ds} = \begin{cases} \lambda e^{-\lambda s} & \text{for } 0 \leqslant s < \infty, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Thus, the delay behaves exponentially with mean $1/\lambda$ and variance $1/\lambda^2$ for the *uniform mobility model*. We conclude from Eqs. (11)–(13) that the average packet delivery delay in the destination cell is $\Theta(n)$ and its variance is $\Theta(n^2)$, i.e.,

$$E[S] = \frac{1}{\lambda} = \Theta(n) \quad \text{and} \quad \text{Var}[S] = \frac{1}{\lambda^2} = \Theta(n^2). \tag{14}$$

Also, from Eqs. (12) and (13), we have that, the delay values are not bounded as a consequence of the tail of the exponential distribution even if the number of total nodes in the network n is finite! Thus, the packet delivery time in the destination cell can last to infinity for some packets, even though its average value is limited by Eq. (14) and n is finite.

5.2. Multi-copy forwarding case

Now consider that K-copies of the same packet were successfully received by adjacent nodes in the destination cell. Obviously, $K \ll n$ as the distribution of the nodes in each cell is assumed to be uniform and we might expect only a small number of nodes within r_0 from the sender. Let $P_D(s)$ be the probability of having the first (and only) delivery of the packet at time interval length s. Hence, given that only one-copy delivery is enforced, and all K relays are looking for the destination, we have that

$$P_D(s) = P\left\{\bigcup_{i=1}^{K}[|X_i(s) - X_{\text{dest}}(s)| \leqslant r_0|s]\right\}. \tag{15}$$

Using union bound and considering that $P_D(s)$ can be at most equal to 1, we arrive at

$$P_D(s) \leqslant \min\left[\sum_{i=1}^{K} P\{|X_i(s) - X_{\text{dest}}(s)| \leqslant r_0|s\}, 1\right], \tag{16}$$

in which Eq. (12) holds for each individual relay i because all the K nodes have independent and identically distributed movements and one can use the results in [4] for a single relay. However, when we attempt to compute the probabilities of multiple relays, since all these nodes start moving from the same area to search for destination (within a circle of radius r_0), their probability distributions are not mutually exclusive. If the time necessary for all these nodes to uniformly spread in the destination cell is equal to t_{spread}, since each node has a speed $v = \Theta(\frac{1}{\sqrt{n}})$, then in general, $t_{\text{spread}} = \Theta(\sqrt{n})$. However, as we will show later, the maximum delay $d_K^{\max} = \Theta(n)$ given that $K \ll n$ *whp*. Therefore, $t_{\text{spread}} \ll d_K^{\max}$ for large values of n, and consequently we can approximate all K probabilities using Eq. (12). This approximation for Eq. (16) results in

$$P_D(s) \leqslant \min[K \cdot P\{|X_1(s) - X_{\text{dest}}(s)| \leqslant r_0|s\}, 1]. \tag{17}$$

Furthermore, Eq. (17) describes two cases. The first case is when $P_D(s)$ is less than 1 while the second case is when the union bound is greater than 1. Obviously, we can derive a meaningful description for d_K only for the first case and that is the basis of our remaining analysis. From Eqs. (12) and (17) and changing s by d_K[3] to indicate the delay for K-copies forwarded to the destination cell, we have for the *uniform mobility model*,

$$\begin{aligned} E_U[P_D(s)] &= E_U\left[P\left\{\bigcup_{i=1}^{K}[|X_i(s) - X_{\text{dest}}(s)| \leqslant r_0|s = d_K]\right\}\right] \\ &= P\{D_K \leqslant d_K\} = F_{D_K}(d_K) \approx K(1 - e^{-\lambda d_K}), \end{aligned} \tag{18}$$

[3] To be more accurate, we should use $\tilde{d}_K$ instead of d_K for the rest of the paper because of the approximation. In order to make the paper easy to read, we will continue to use the same notation for simplicity.

for a uniform steady-state distribution resulting from the random motion of the nodes. Exact computation of probability of d_K is a tedious task, instead, we assume that the upper bound probability can be achieved while this is simply an approximation. We make this assumption to find an approximate description for d_K and then by using computer simulation for MANETs, given in the next subsection and in [8,9], we showed that this approximation can demonstrate the asymptotic behavior of d_K reasonably well (see Fig. 4). $F_{D_K}(d_K)$ can be interpreted as the cumulative density function of the delay random variable D_K for K-copies transmission to the destination cell.

From Eq. (18) we see that the maximum value attained by D_K is given when

$$F_{D_K}(d_K^{\max}) = 1 \approx K(1 - e^{-\lambda d_K^{\max}}) \Rightarrow d_K^{\max} \approx \frac{1}{\lambda} \log\left(\frac{K}{K-1}\right). \quad (19)$$

Eq. (19) reveals that, *for a finite n*, the new delay obtained by multi-copy forwarding is bounded by $d_K^{\max}$ after ensemble averaging over all possible starting points that are uniformly distributed (see Fig. 4).

From Eqs. (11) and (19), and because $K \ll n$ *whp*, $d_K^{\max}$ grows to infinity if n scales to infinity.

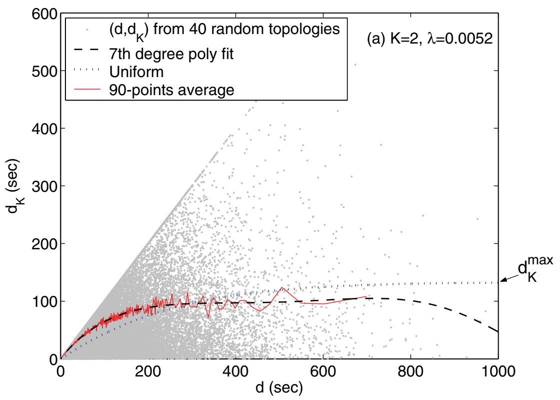

Fig. 4. Simulation results for the *random waypoint mobility model*. Each grey point is a pair (d, d_K) delay measured for 40 random topologies all plotted together. A 7th degree polynomial fit for all the points and a 90 consecutive points average are plotted for $K = 2$. The theoretical curve for the steady-state uniform distribution is also plotted.

The probability density function for D_K is

$$f_{D_K}(d_K) = \frac{dF_{D_K}}{dd_K} \approx \begin{cases} K\lambda e^{-\lambda d_K} & \text{for } 0 \leqslant d_K \leqslant d_K^{\max}, \\ 0 & \text{otherwise.} \end{cases} \quad (20)$$

Hence, in the multi-copy forwarding scheme ($K > 1$) the tail of the exponential delay distribution is cut off. Note that the time-to-live threshold in the destination cell must be set greater than the worst asymptotic delay ($K = 2$) to allow the packet to be delivered, i.e., $d_2^{\max} \approx \frac{\log(2)}{\lambda} < \mathrm{TTL}$.

As in *Scheme 2*, the total delivery delay for a packet, measured from the source to the destination, is divided in two parts: the time the packet spends to reach the destination cell, plus the time the relay in the destination cell expends to reach the destination node. The former was shown to be $\Theta(\sqrt{n})$, and for a fixed n this delay is finite. However, as discussed above, the latter can last indefinitely if only one copy is looking for the destination. Hence, by forwarding K-copies in the destination cell, the total delivery delay is approximated by

$$D_{2_K} \approx \Theta(\sqrt{n}) + d_K. \quad (21)$$

Thus, a delay of hours in single-copy forwarding to the destination cell can be reduced to a few minutes or even a few seconds for multi-copy relaying, depending on the network parameter values.

We have shown [8,9] that the throughput per source–destination pair for the multi-copy relaying approach remains at $\Theta(1)$ [2]. Thus, by multi-copy forwarding at the destination cell in the modified version of *Scheme 2*, we do not change the order of the capacity. Hence, Theorem 2 still holds here.

5.3. Simulation results

To compare the approximation for delay analysis, we have simulated our multi-copy forwarding strategy. The *BonnMotion* simulator [13] was used, which creates mobility scenarios that can be utilized to study mobile ad hoc network characteristics.

We implemented the *random waypoint mobility model* [14,15] for the random motion of the nodes (as it resembles the *uniform mobility model* [5]). In this model, nodes are initially randomly distributed in the network area. The model is also characterized by the time a node remains in one position, called pause time, and the movement of a node toward a random destination point in the network with a speed range between $v_{\min}$ and $v_{\max}$. The nodes movement are independent of each other. In our simulation, the pause time is zero, and $v_{\min} = v_{\max} = v$.

Fig. 4 shows the results for 1000 s of simulations for $n = 1000$ nodes, $v = 0.13$ m/s, $r_0 = 0.02$ m, and a unit area disk as the simulation area, which results $\lambda = 0.0052$. In this case, the simulation area is considered as the destination cell with n nodes. To obtain a solution close to the steady-state behavior, we ran 40 random topologies and averaged them as follows. In each run we chose randomly a node with $K = 2$ neighbors, within r_0, and measured the time that each of these K nodes reached each of the other $n - K$ nodes in the disk (i.e., except the sender and its other $K - 1$ neighbors) considering each of them as a destination. The delay of the sender's nearest node reaching each destination is by definition d, and d_K is the minimum time among all the K nodes that reach the destination. Fig. 4 shows all pairs of points (d, d_K) obtained in this way for $K = 2$. In each graph we plot a 7th degree polynomial fit for all the points as well as an average obtained by taking the mean of consecutive 90 points. We also plot the theoretical curve (from Eq. (40) in [9]) for the steady-state uniform distribution for the same parameters. We see that the averaged 90-points curve follows the polynomial fit and that they both accompany the steady-state uniform distribution predicted by theory as they are related mobility models. We only observe the asymptotic behavior for the experimental curves up to 800 s. The polynomial fit begins to fall after that, and does not represent the actual asymptotic behavior anymore due to the natural lack of samples in this part of the graph.

6. Fixed nodes with directional antennas

In this section, we present a model where nodes are static, but endowed with directional antennas. Previous works [10,16] have considered capacity analysis for static networks using directional antennas, where they showed that no scheme using directed beams can circumvent the constriction on capacity in dense networks. In our study, we present a slightly different modeling approach compared to these previous directional antenna analysis. We constrain communication to occur only between closest neighbors by using very narrow beams. The network model is shown in Fig. 5. A source–destination pair of nodes is randomly chosen so that we want to send a packet from cell a to cell t, for example, relying on multiple relays (or hops) using directional antenna transmission along close neighbors in the path to the destination. The nodes are deployed uniformly in the network area torus. As in *Scheme 1*, the network is divided in $1/a(n)$ cells, each with an area $a(n)$. We assume $a(n) \geqslant 2\log(n)/n$, so that each cell has at least one node *whp* [7]. In each cell a node is chosen to relay the traffic of the cell. Fig. 5 shows a source node in cell a that has destination at a node in cell t separated by a distance $\overline{L}$. Accordingly, the cell path along the closest neighbors is $\{a \to f \to g \to h \to m \to n \to o \to t\}$.

We want to obtain the average throughput for a source–destination pair uniformly chosen among all n nodes, as well as the delay behavior. The relay transmissions are scheduled at regular time intervals

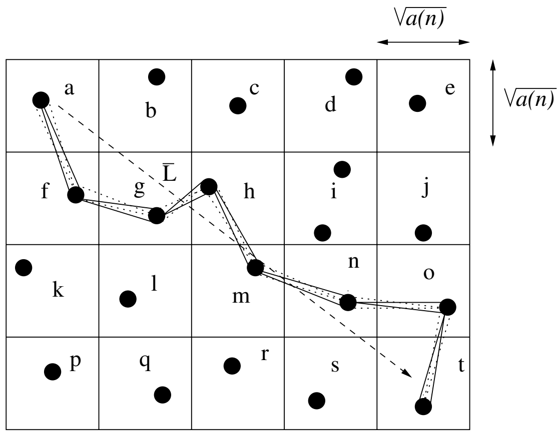

Fig. 5. Unit area torus network divided into $1/a(n)$ cells each with size area of $a(n)$. Transmissions are employed using bi-directional antennas, with very narrow beams, between closest neighbors from adjacent cells along the path to destination.

so that each node is assigned a time slot to transmit successfully to its closest neighbor in the path to the chosen destination. This is a time schedule constraint because a node can only point its antenna to a close neighbor at consecutive time intervals. For the example shown in Fig. 5, each node has eight neighbors, given that we assume a torus net, so that it can communicate to each of them at regular eight slot time interval respectively, i.e., a time division multiple access (TDMA) with bi-directional beam transmission. Each time two nodes point their antennas to each other, they exchange packets, so that each of these exchanges can involve either source–relay, relay–relay, or relay–destination transmissions. Interference is overcome by the use of directional beams to the nearest neighbor, so that Eq. (1) is satisfied. Again we assume that the transmissions are half-duplex, i.e., the communication time slot is divided in two equal parts. Each node transmit at W bits/s. So the average available bit rate is $W/2$ bits/s.

Given that $\overline{L}$ is the mean distance between two uniformly and independently chosen source–destination pair in the network, the average path distance across cells traversed by a packet is $\Theta(\overline{L})$. Accordingly, each cell hop has average size of $\sqrt{a(n)}$. Thus, the mean number of hops traversed by a packet until it reaches its destination is $\frac{\Theta(\overline{L})}{\sqrt{a(n)}}$.

According to the definition of throughput, each source generates $\Lambda(n)$ bits/s. Given that each bit needs to be relayed on the average by $\frac{\Theta(\overline{L})}{\sqrt{a(n)}}$ nodes, the total average number of bits per second served by the entire network equals $\frac{\Theta(\overline{L})n\Lambda(n)}{\sqrt{a(n)}}$. To ensure that all required traffic is carried, we need that

$$c_{12}n\frac{W}{2}\Delta t \leqslant \frac{\Theta(\overline{L})n\Lambda(n)}{\sqrt{a(n)}} \leqslant c_{13}n\frac{W}{2}\Delta t, \tag{22}$$

where $\Delta t = 1/8$, which comes from the TDMA transmission schedule approach.[4] Thus,

$$c_{14}\sqrt{a(n)} \leqslant \Lambda(n) \leqslant c_{15}\sqrt{a(n)}. \tag{23}$$

This proves the following theorem.

[4] Other diversity scheme could be assumed as well.

Theorem 3. *For a given node using directional antenna transmission to closest neighbor along the path to destination, with* $a(n) = \frac{k\log(n)}{n}$, *for* $k \geqslant 2$, *to guarantee connectivity, we have*

$$\Lambda_D(n) = \Theta\left(\sqrt{\frac{\log(n)}{n}}\right).$$

This result represents a better bound on throughput capacity than what Gupta and Kumar [1] obtained which was $\Theta(1/\sqrt{n\log(n)})$, and the results by Yi et al. [10]. Indeed, it is a gain of $\Theta(\log(n))$ and is similar to Peraki and Servetto's results [16] obtained for a single directed beam, where they use a different approach applying networking flow analysis to calculate the network transport capacity (i.e., maximum stable throughput). This is the same capacity scalability obtained for *Scheme 1*. We see that capacity is still constrained in dense networks. It is due to the wasting of the available bandwidth to forward the same packet over multiple hops by an amount of time that scales with n.

The average delay incurred by a packet to reach the destination is the sum of the average time a packet spends hopping along the path to its destination. The total number of hops to reach destination is $\Theta(\overline{L}/\sqrt{a(n)})$. Accordingly, the delay using directional antenna transmission to nearest neighbor is given by

$$D_D(n) = (\#\text{ of hops})\Delta t = \Theta\left(\frac{1}{\sqrt{a(n)}}\right) = \Theta\left(\sqrt{\frac{n}{\log(n)}}\right). \tag{24}$$

Compared to Eq. (6) this represents a delay reduction of $\Theta(1/\sqrt{\log(n)})$. Thus, the use of directional antenna with fixed nodes offers a smaller delay on average than the restricted mobility case, while attaining the same throughput scalability as *Scheme 1*.

Therefore, employing directional antenna transmissions between closest nodes along the path to a destination is equivalent, in terms of throughput performance, to nodes executing restricted mobility as in *Scheme 1*, while providing a smaller packet delivery delay.

Table 1
Throughput gain and delay increase obtained from comparing previous works [1,2] with restricted mobility schemes and directional antenna transmission

Schemes comparisons	Throughput gain	Delay increase
$\frac{\text{Scheme 1}}{\text{Gupta and Kumar}}$	$\log(n)$	$\sqrt{\log(n)}$
$\frac{\text{Grossglauser and Tse}}{\text{Scheme 2}}$	$\sqrt{l}$	l
$\frac{\text{Directional antenna}}{\text{Gupta and Kumar}}$	$\log(n)$	None
$\frac{\text{Scheme 1}}{\text{Directional antenna}}$	None	$\sqrt{\log(n)}$

7. Performance comparisons

To obtain a benchmark of throughput and delay for wireless ad hoc networks, we compare in Table 1 the schemes studied with the previous works by Gupta and Kumar [1], and Grossglauser and Tse [2]. The results suggest that using mobility or enhanced physical layer properties (directional antennas in this case) can improve throughput or delay.

8. Conclusions

We have analyzed four schemes for ad hoc wireless networks. The first three schemes considered nodes with restricted mobility. The nodes have restrained mobility area that can be either a function of n, or independent of n. We show that on all these cases we can trade-off the mobility resource with capacity and delay. In the first scheme the capacity does not scale well, while in the second scheme the throughput has non-zero asymptotic behavior in dense networks, and it is shown to be a generalization of the Grossglauser and Tse [2] results. The third scheme is a modified version of the second, in which we allow multiple packet copies to be forwarded to the destination cell so that we attain a better delay performance. The fourth scheme studied was that of a static ad hoc network using directional antennas with transmission restricted to closest neighbors in the path along destination. We showed that the capacity still decreases with n having the same scalability law as that obtained in the first scheme of restricted mobility, however presenting a smaller delay. Therefore, the directional antenna scheme provides better throughput performance than static networks employing omnidirectional antennas, and presents smaller delay than in restricted mobility.

Acknowledgments

This work was supported in part by CAPES/Brazil, the UCSC Grant 19900-444068-EEFRHS, the Baskin Chair of Computer Engineering, by DARPA under grant W15P7T-05-C-P015, and by the Army Research Office under grant W911NF-04-1-0224.

The authors would like to thank the reviewers for their constructive comments and suggestions.

References

[1] P. Gupta, P.R. Kumar, The capacity of wireless networks, IEEE Transactions on Information Theory 46 (2) (2000) 388–404.
[2] M. Grossglauser, D. Tse, Mobility increases the capacity of wireless ad hoc networks, IEEE/ACM Transactions on Networking 10 (4) (2002) 477–486.
[3] S.N. Diggavi, M. Grossgaluser, D. Tse, Even one-dimensional mobility increases ad hoc wireless capacity, in: Proc. IEEE ISIT, Laussane, Switzerland, 2002.
[4] E. Perevalov, R. Blum, Delay limited capacity of ad hoc networks: asymptotically optimal transmission and relaying strategy, in: Proc. IEEE Infocom, San Francisco, CA, 2003.
[5] N. Bansal, Z. Liu, Capacity, delay and mobility in wireless ad-hoc networks, in: Proc. IEEE Infocom, San Francisco, CA, 2003.
[6] M.J. Neely, E. Modiano, Improving delay in ad-hoc mobile networks via redundant packet transfers, in: Proc. Conference on Information Sciences and Systems, Baltimore, MD, 2003.
[7] A.E. Gamal, J. Mammen, B. Prabhakar, D. Shah, Throughput-delay trade-off in wireless networks, in: Proc. IEEE Infocom, Hong Kong, 2004.
[8] R.M. de Moraes, H.R. Sadjadpour, J.J. Garcia-Luna-Aceves, Making ad hoc networks scale using mobility and multi-copy forwarding, in: Proc. of IEEE Globecom, Dallas, TX, 2004.
[9] R.M. de Moraes, H.R. Sadjadpour, J.J. Garcia-Luna-Aceves, Throughput-delay analysis of mobile ad-hoc networks with a multi-copy relaying strategy, in: Proc. IEEE SECON, Santa Clara, CA, 2004.

[10] S. Yi, Y. Pei, S. Kalyanaraman, On the capacity improvement of ad hoc wireless networks using directional antennas, in: Proc. ACM MobiHoc, Annapolis, MD, 2003.
[11] M. Franceschetti, O. Dousse, D. Tse, P. Tiran, Closing the gap in the capacity of random wireless networks, in: Proc. IEEE ISIT, Chicago, IL, 2004.
[12] R. Knopp, P.A. Humblet, Information capacity and power control in single-cell multiuser communications, in: Proc. IEEE ICC, Seattle, WA, 1995.
[13] C. de Waal, M. Gerharz, Bonnmotion: a mobility scenario generation and analysis tool. Available from <http://www.cs.uni-bonn.de/IV/BonnMotion/>.
[14] T. Camp, J. Boleng, V. Davies, A survey of mobility models for ad hoc network research, in: Mobile Ad Hoc Networking: Research, Trends and Applications, Wireless Communication & Mobile Computing 2 (5) (2002) 483–502 (special issue).
[15] C. Bettstetter, H. Hartenstein, X. Pérez-Costa, Stochastic properties of the random waypoint mobility model: epoch length, direction distribution, and cell change rate, in: Proc. ACM MSWiM, Atlanta, GA, 2002.
[16] C. Peraki, S.D. Servetto, On the maximum stable throughput problem in random networks with directional antennas, in: Proc. ACM MobiHoc, Annapolis, MD, 2003.

Renato M. de Moraes received the electrical engineering degree from Federal University of Pernambuco—Brazil in 1996, and the Master degree from UNICAMP—Sao Paulo—Brazil in 1998. Currently, He is in the Department of Electrical Engineering at University of California, Santa Cruz, where he is working toward the Ph.D. degree. His areas of interest include Ad hoc Networks, Information Theory and Coding, and Communication Networks.

Hamid R. Sadjadpour received his B.S. and M.S. degrees from Sharif University of Technology with high honor and Ph.D. degree from University of Southern of California in 1986, 1988, and 1996 respectively. After graduation, he joined AT&T as a Member of technical staff, later Senior technical staff member, and finally Principal member of technical staff at AT&T Research Lab. in Florham Park, NJ until 2001. In fall 2001, he joined University of California, Santa Cruz (UCSC) as an Assistant Professor. He has served as technical program committee member in ICC 2003, RAWCON 2004, Globecom2004, ICC 2005, RAWCON 2006, Globecom 2005, ICC 2006, Guest Editor of EURASIP on special issue on Multicarrier Communications and Signal Processing, 2003, and Chair of Communication Theory Symposium at WirelessCom 2005. He has published more than 60 journal, conference, book chapter, and technical reports. His research interests include communication theory and signal processing, coding theory for wireless, wired, and fiber optic applications, space-time signal processing for wireless systems, Turbo codes, and Information theory for ad hoc networks. He holds more than 13 patents.

J.J. Garcia-Luna-Aceves received the B.S. degree in electrical engineering from the Universidad Iberoamericana in Mexico City, Mexico in 1977, and the M.S. and Ph.D. degrees in electrical engineering from the University of Hawaii, Honolulu, HI, in 1980 and 1983, respectively. He holds the Jack Baskin Chair of Computer Engineering at the University of California, Santa Cruz (UCSC). He is also a Principal Scientist at the Palo Alto Research Center (PARC). Prior to joining UCSC in 1993, he was a Center Director at SRI International (SRI) in Menlo Park, California. He has been a Visiting Professor at Sun Laboratories and a Principal of Protocol Design at Nokia.

He has published a book, more than 290 papers, and seven US patents. He has directed 21 Ph.D. theses and 19 M.S. theses at UCSC over the past 11 years. He is the General Chair for the IEEE SECON 2005 Conference. He has also been Program Co-Chair of ACM MobiHoc 2002 and ACM Mobicom 2000; Chair of the ACM SIG Multimedia; General Chair of ACM Multimedia '93 and ACM SIGCOMM '88; and Program Chair of IEEE MULTIMEDIA '92, ACM SIGCOMM '87, and ACM SIGCOMM '86. He has served in the IEEE Internet Technology Award Committee, the IEEE Richard W. Hamming Medal Committee, and the National Research Council Panel on Digitization and Communications Science of the Army Research Laboratory Technical Assessment Board. He has been on the editorial boards of the IEEE/ACM Transactions on Networking, the Multimedia Systems Journal, and the Journal of High Speed Networks. He received the SRI International Exceptional-Achievement Award in 1985 and 1989, and is a senior member of the IEEE.