On Mobility-Capacity-Delay Trade-off in Wireless Ad Hoc Networks *

Renato M. de Moraes, Hamid R. Sadjadpour
Department of Electrical Engineering
University of California at Santa Cruz
Santa Cruz, CA 95064, USA
Email: {renato,hamid}@soe.ucsc.edu

J.J. Garcia-Luna-Aceves
Department of Computer Engineering
University of California at Santa Cruz
Santa Cruz, CA 95064, USA
Email: jj@soe.ucsc.edu

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 mobility of nodes 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. By restricting the movement of nodes within these cells, we 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.

1. Introduction

Capacity analysis in ad hoc networks has become an important issue since Gupta and Kumar [6] showed that the capacity of a fixed and connected wireless network decreases as the number of nodes n increases. Grossglauser and Tse [5] presented a two-phase packet forwarding technique for mobile ad hoc networks (MANET) 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 [5] 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 have traded off delay in order to attain higher capacity in mobile ad hoc networks (MANETs) [5], [3], [10], [1], [8], [4].

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. As nodes have restrained mobility, the delivery from source to destination is done across multiple hops obtained by relaying this packet 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.

Before summarizing the main results of the paper we review the following definitions [6], [4]:

A *throughput* of $\Lambda(n)$ bits per second is feasible if every node can send information at a rate of $\Lambda(n)$ bits per second to its chosen destination.

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 known network model that have been used recently to analyze the capacity of wireless network [4], [6], [5], [1]. Section 3 presents a restricted mobility model, called *Scheme 1*, where the size of the cells varies with the number of nodes n. The as-

* This work was supported in part by CAPES/Brazil, a UCSC COR Grant, and the Baskin Chair of Computer Engineering.

sociated throughput and delay are given by[1]

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

Compared to the static network model [6], *Scheme 1* attains a gain of $O(log(n))$ by using restrained mobility.

Section 4 presents another restricted mobility model, called *Scheme 2*, in which the size of a cell is not a function of n. Indeed, for a given constant number of cells l, the size of each cell is $1/l$, and the corresponding throughput and delay are

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

This throughput result is a generalization of the results by Grossglauser and Tse [5] 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 the *Scheme 2* to allow multiple-copy relaying [2] so that the order of magnitude of the throughput is preserved, but bounded delivery delay is attained when the numbers of total nodes (n) is finite.

Section 6 presents the throughput-delay analysis for a fixed network where nodes are endowed with directional antennas. Nodes relay packets to their closest neighbors along the path to destinations. We find that

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

This result is important, because it represents a capacity gain of $O(log(n))$ compared to the results in Gupta and Kumar [6], and Yi et al [11].

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 [4], [6], [5].

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/sec* to j if [6]

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

where $\Delta > 0$, so that node X_k will not impede X_i and X_j communication. This is called the *protocol model* [6].

For the case of *mobile nodes*, the position of node i at any time is now a function of time. The nodes are assumed to be uniformly distributed on the torus and there is no preferential direction of movement where each node moves with speed $v(n)$. The trajectories for different nodes are independent and identically distributed. A successful transmission between nodes i and j is governed again by Eq. (1), where the position of the nodes are time dependent [5]. Time is slotted to simplify the analysis. Also, at each time step, a scheduler decides which nodes are senders, 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* [1]. 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 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 [1]: (a) At any time t, the position of the nodes are independent of each other; (b) the steady-state distribution of the mobile nodes is uniform; and (c) conditional on the position of a node, the direction of the node movement is uniformly distributed in $[0, 2\pi)$.

3. Scheme 1

We present a restricted mobility scheme that attains a capacity gain of $log(n)$ compared to the static network model [6]. The throughput still 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 shown in Fig. 1. The network is a unit torus divided in square cells, each of area $a(n)$ as in [4], and they showed that, if $a(n) \geq \frac{2\,log(n)}{n}$, then each cell has at least one node *with high probability (whp)*, i.e., with probability $\geq 1 - 1/n$. This condition guarantees connectivity *whp* [6], [4].

We now consider an 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* that moves with speed $v(n)$, and has no preferential direction of movement within the cell. The nodes move independently of each

1 Here we use the Knuth's notation: (a) $f(n) = O(g(n))$ means there are positive constants b_1 and N_1, such that $0 \leq f(n) \leq b_1 g(n)\ \forall\, n \geq N_1$. (b) $f(n) = \Theta(g(n))$ means there are positive constants b_2, b_3, and N_2, such that $0 \leq b_2 g(n) \leq f(n) \leq b_3 g(n)\ \forall\, n \geq N_2$.

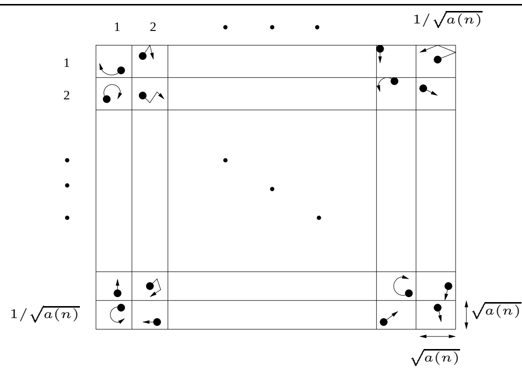

Figure 1. Unit area torus network divided into $1/a(n)$ cells, each with size area of $a(n)$.

other, and once they hit the cell boundaries they are bounced back (with relation to the normal edge).

We assume that each node only communicates with another node from any adjacent cell, and this happens only when they are close enough to each other (i.e., both are near to the common edge that separates the cells) so that the effect of interference can be minimized. Thus, a source node will rely on relays across several cells to have its 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 shows a packet whose source node is in cell i and has as its destination a node in cell d, separated by an average distance $\overline{L}$. Possible cell paths for this packet are $\{ijfgcd\}$, $\{ijfghd\}$, $\{iefgcd\}$, $\{iefghd\}$, for example.

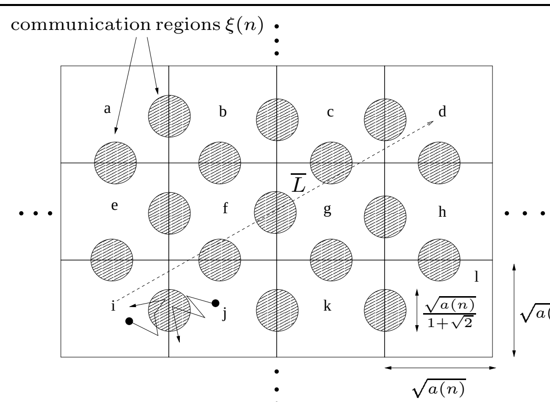

Figure 2. Region $\xi(n)$ where communication between nodes from adjacent cells is possible.

Grossglauser and Tse [5] 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 rate of W *bits/sec*, and the other half to receive at also W *bits/sec*. Thus, the average available bandwidth is $\frac{W}{2}$ *bits/sec*. At each time two nodes communicate with each other, they exchange packets, so that each of these sessions can be source-relay, relay-relay, or relay-destination transmission.

The area where successful communication can occur is shown in Fig. 2. Basically, it is a semi-circumference $\xi(n)$ 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 $\xi(n)$, 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 $\xi(n)$'s in their cells as shown. We have that

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

Now, the probability of finding a node traveling inside $\xi(n)$ is $\frac{\xi(n)}{a(n)}$, because the node has no preferential direction of movement in the cell and tends to move uniformly inside the cell. In addition, because the nodes have independent and identically distributed (iid) movements, the probability that both nodes come to the communication region simultaneously, denoted by P_c, equals

$$P_c = \left[\frac{\xi(n)}{a(n)}\right]^2 = \left(\frac{\pi}{24+16\sqrt{2}}\right)^2 = c_1. \tag{3}$$

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 $O(\overline{L})$. Accordingly, each cell hop has an average size of $\sqrt{a(n)}$. Thus, the mean number of hops traversed by a packet is $\frac{O(\overline{L})}{\sqrt{a(n)}}$.

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

$$\frac{O(\overline{L})n\Lambda(n)}{\sqrt{a(n)}} \leq n\,\frac{W}{2}P_c \Longrightarrow \Lambda(n) \leq c_2 W\sqrt{a(n)}. \quad (4)$$

We just proved the following Theorem.

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

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

Compared to the capacity result obtained in [6], the result of Theorem 1 represents a gain of $O(log(n))$. *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 multiuser diversity [7] to overcome interference (note that Gupta and Kumar [6] could not use multiuser diversity because they consider only fixed nodes), the network still does not scale well with the number of nodes, i.e., $\Lambda_1(n) \longrightarrow 0$ 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 [5].

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

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

where $\Delta S = O(\sqrt{a(n)})$ is the average distance in one-round trip inside a cell. Note also that the total number of hops is $O(\overline{L}/\sqrt{a(n)})$, 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 where 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 [4]. Accordingly, 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

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

This delay is larger than that obtained by Gupta and Kumar [6], which was shown to be $O(1/\sqrt{a(n)}) = O(\sqrt{n/log(n)})$ [4]. This is a direct consequence of the throughput-delay trade-off property [4]. *The improvement of capacity 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 [5].

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 $\frac{1}{l}$. 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 [5]. 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 normal edge. Here, we consider that each node can communicate with its closest neighbor within the transmission range r_o, 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, $r_o = 1/\sqrt{\theta\pi n}$ [2]. 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.

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 will

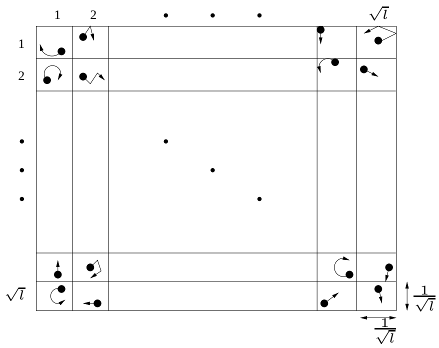

Figure 3. Unit area torus network divided into l cells, each with size area of $\frac{1}{l}$.

not be relayed again for another node in the same cell. Instead, the node will keep the packet in its queue until it reaches the neighborhood of an adjacent cell in the path toward to 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_o, then it relays the packet to this neighbor if there is a packet to forward in that direction, so that it can be either source-relay, or relay-relay, or relay-destination transmission. The communication is simplex so that each node uses the entire communication time slot to transmit at rate W bits/sec.

From the above description, we claim that in *steady-state*:

- *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 $2log(n)/n$), and because of the uniform distribution of the nodes, each cell will contain $O(\frac{n}{l})$ nodes. Since $n \longrightarrow \infty$, l can be any positive integer and is not a function of n.

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

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

$$\frac{O(\overline{L})n_S\Lambda(n)}{1/\sqrt{l}} \leq n_S\,W \Longrightarrow \Lambda(n) \leq \frac{c_3 W}{\sqrt{l}}. \qquad (7)$$

This proves the following Theorem.

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

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

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

Because a packet is not allowed to move through the entire network (waiting in the relay queue as in [5]), it follows 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 [5]. 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 expends to reach the destination node. The later is $O\left(\frac{n}{l}\right)$ as we have $O\left(\frac{n}{l}\right)$ nodes in each cell [4], [8]. The former is given by the number of hops traversed multiplied by the time spent per hop (i.e., (# of hops) $\cdot\, t(n)$) which is $O\left[\left(\frac{\overline{L}}{1/\sqrt{l}}\,\frac{1}{v(n)}\right)\,\frac{1}{\sqrt{l}}\right] = O(\sqrt{n})$. However, for a sufficiently large value of n (and $l << \sqrt{n}$), the term $\frac{n}{l}$ dominates $\sqrt{n}$, and

$$D_2(n) = O\left(\frac{n}{l}\right). \qquad (8)$$

Comparing $D_2(n)$ to the delay attained in the scheme by Grossglauser and Tse [5], whose delay was shown to be $O(n)$ [8], [4], 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 [5], 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 $\frac{1}{l}$, the $O(1)$ throughput obtained in [5] 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 [5].

Although the average delivery delay is given by Eq. (8), we have shown [2] that the delivery delay is an exponential random variable for a mobile ad hoc network.

Accordingly, from the tail of the exponential distribution, this delay can last to infinity even when n is finite [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 finite delivery delay is possible for finite values of n.

5. Scheme 2 with Multi-Copy Relaying at Destination Cell

We now introduce an improved packet forwarding strategy [2] for mobile ad hoc networks that attains the $\Theta(1)$ capacity of the basic scheme by Grossglauser and Tse [5], but provides bounded delay when the number of nodes n is fixed.

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 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_o 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 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 $K = 1$, it has been shown [2], [10] that the delivery delay random variable d has an exponential distribution with parameter $\lambda = 2\, r_o\, v$ which results from evaluating the flux of nodes entering a circle of radius r_o during a differential time interval considering the nodes uniformly distributed over the entire network of unit area and traveling at speed v. Accordingly, even for finite n, the delivery delay can last to infinity. For a uniform distribution of the nodes, $r_o = 1/\sqrt{\pi\theta n}$. Hence, the radius r_o decreases with $1/\sqrt{n}$. The velocity of the nodes also decreases with $1/\sqrt{n}$. Hence, $\lambda = \frac{1}{\Theta(n)}$. In [2] we extend this model to consider the case $K > 1$ and find that the tail of the exponential distribution is cut off resulting a new delivery delay random variable d_K which is related to d by [2]

$$d_K \approx \frac{1}{\lambda}\, ln\left(\frac{K}{K-1+e^{-\lambda d}}\right), \tag{9}$$

for a steady-state uniform distribution resulting as motion of nodes. Hence, for a fixed n and $1 < K << n$, the maximum delay d_K tends to a constant value as d increases. Thus, the delivery delay is bounded for finite n.

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 $O(\sqrt{n})$, and for a fixed n this delay is finite. However, the latter can last to infinity as discussed above if only one copy is looking for the destination. Hence, by forwarding K-copies in the destination cell, the total delivery delay is given by

$$D_{2_K} = O(\sqrt{n}) + d_K, \tag{10}$$

that is finite for a fixed n. 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 [2] that the throughput per source-destination pair for the multi-copy relaying approach remains at $\Theta(1)$ [5]. 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.

6. Fixed Nodes with Directional Antennas

In this section, we present a model where nodes are static, but endowed with directional antennas. Previous works [11], [9] 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 slight different modeling approach compared to previous directional antenna analysis [11], [9], where we constrain the communication only between closest neighbors by using very narrow beams. The network model is shown in Fig. 4. 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 $a(n) \geq 2\,log(n)/n$, so that each cell has at least one node *whp* [4]. Fig. 4 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 $\{afghmnot\}$.

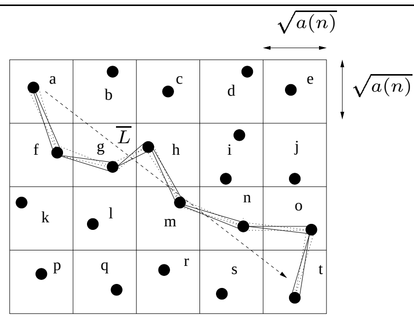

Figure 4. Unit area torus network divided into $1/a(n)$ cells each with size area of $a(n)$. Transmissions are employed using bi-directional antennas, with very narrow beams, between closest neighbors from adjacent cells along the path to destination.

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 so that each node is assigned a time slot to transmit successfully to its closest neighbor in the path to destination. This is a time schedule constraint as a node can only point its antenna to a close neighbor at consecutive time intervals. For the example shown in Fig. 4, 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. At each time that two nodes point their antennas to each other they exchange packets, so that each of these sessions can be either source-relay, relay-relay, or relay-destination transmission. 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/sec*. So the average available bandwidth is $W/2$ *bits/sec*.

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 $O(\overline{L})$. Accordingly, each cell hop has average size of $\sqrt{a(n)}$. Thus, the mean number of hops traversed by a packet until destination is $\frac{O(\overline{L})}{\sqrt{a(n)}}$.

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

$$\frac{O(\overline{L})n\Lambda(n)}{\sqrt{a(n)}} \leq n\,\frac{W}{2}\Delta t, \qquad (11)$$

where $\Delta t = \frac{1}{8}$, which comes from the TDMA transmission schedule approach[2]. Thus,

$$\Lambda(n) \leq c_4 W\sqrt{a(n)}. \qquad (12)$$

This proves the following Theorem.

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* $k \geq 2$, *to guarantee connectivity, we have*

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

This result represents a better bound on throughput capacity than the $O(1/\sqrt{n\,log(n)})$ obtained by Gupta and Kumar [6], and those obtained in [11]. Indeed, it is a gain of $O(log(n))$ and agrees with Peraki and Servetto's results [9] 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

2 Other diversity scheme could be assumed as well.

$O(\overline{L}/\sqrt{a(n)})$. Accordingly, the delay using directional antenna transmission to nearest neighbor is given by

$$D_D(n) = (\#\text{ of hops})\Delta t = O\left(\frac{1}{\sqrt{a(n)}}\right) = O\left(\sqrt{\frac{n}{log(n)}}\right). \quad (13)$$

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

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

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 [6], and Grossglauser and Tse [5].

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

Table 1. Throughput gain and delay increase obtained from comparing previous works [6], [5] with restricted mobility schemes and directional antenna transmission.

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 [5] results. The third scheme is a modified version of the second, in which we allow multiple packet copies to be forwarded at the destination cell so that we attain bounded delay for a finite number of nodes n. 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.

References

[1] N. Bansal and Z. Liu. Capacity, delay and mobility in wireless ad-hoc networks. In *Proc. of IEEE Infocom*, March 2003.

[2] R. M. de Moraes, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves. Throughput-delay analysis of mobile ad-hoc networks with a multi-copy relaying strategy. In *Proc. of IEEE SECON*, October 2004.

[3] S. N. Diggavi, M. Grossgaluser, and D. Tse. Even one-dimensional mobility increases ad hoc wireless capacity. In *Proc. of IEEE ISIT*, June 2002.

[4] A. El-Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-delay trade-off in wireless networks. In *Proc. of IEEE Infocom*, March 2004.

[5] M. Grossglauser and D. Tse. Mobility increases the capacity of wireless ad-hoc networks. In *Proc. of IEEE Infocom*, March 2001.

[6] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000.

[7] R. Knopp and P. A. Humblet. Information capacity and power control in single-cell multiuser communications. In *Proc. of IEEE ICC*, June 1995.

[8] M. J. Neely and E. Modiano. Improving delay in ad-hoc mobile networks via redundant packet transfers. In *Proc. of Conference on Information Sciences and Systems*, March 2003.

[9] C. Peraki and S. D. Servetto. On the maximum stable throughput problem in random networks with directional antennas. In *Proc. of ACM MobiHoc*, June 2003.

[10] E. Perevalov and R. Blum. Delay limited capacity of ad hoc networks: asymptotically optimal transmission and relaying strategy. In *Proc. of IEEE Infocom*, March 2003.

[11] S. Yi, Y. Pei, and S. Kalyanaraman. On the capacity improvement of ad hoc wireless networks using directional antennas. In *Proc. of ACM MobiHoc*, June 2003.