Optimal Multicast Capacity and Delay Tradeoffs in MANETs

Jinbei Zhang, Xinbing Wang, *Senior Member, IEEE*, Xiaohua Tian, *Member, IEEE*
Yun Wang, Xiaoyu Chu, and Yu Cheng, *Senior Member, IEEE*

Abstract—In this paper, we give a global perspective of multicast capacity and delay analysis in Mobile Ad Hoc Networks (MANETs). Specifically, we consider four node mobility models: (1) two-dimensional i.i.d. mobility, (2) two-dimensional hybrid random walk, (3) one-dimensional i.i.d. mobility, and (4) one-dimensional hybrid random walk. Two mobility time-scales are investigated in this paper: (i) Fast mobility where node mobility is at the same time-scale as data transmissions; (ii) Slow mobility where node mobility is assumed to occur at a much slower time-scale than data transmissions. Given a delay constraint D, we first characterize the optimal multicast capacity for each of the eight types of mobility models, and then we develop a scheme that can achieve a capacity-delay tradeoff close to the upper bound up to a logarithmic factor. In addition, we also study heterogeneous networks with infrastructure support.

Index Terms—Multicast capacity and delay tradeoffs, Mobile Ad Hoc Networks (MANETs), independent and identically distributed (i.i.d.) mobility models, hybrid random walk mobility models, capacity achieving schemes, heterogeneous networks

---◆---

1 INTRODUCTION

Since the seminal paper by Gupta and Kumar [1], where a maximum per-node throughput of $O(1/\sqrt{n})$ was established in a static network with n nodes, there has been tremendous interest in the networking research community to understand the fundamental achievable capacity in wireless ad hoc networks. How to improve the network performance, in terms of the capacity and delay, has been a central issue.

Many works have been conducted to investigate the improvement by introducing different kinds of mobility into the network, [2], [3], [4], [5], [6], [7]. Other works attempt to improve capacity by introducing base stations as infrastructure support, [8], [9], [10].

As the demand of information sharing increases rapidly, multicast flows are expected to be predominant in many of the emerging applications, such as the order delivery in battlefield networks and wireless video conferences. Related works are [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], including static, mobile and hybrid networks.

Introducing mobility into the multicast traffic pattern, Hu et al. [14] studied a motioncast model. Fast mobility was assumed. Capacity and delay were calculated under two particular algorithms, and the tradeoff derived from them was $\lambda = O(\frac{D}{nk \log k})$, where k was

---

• *Jinbei Zhang, Xinbing Wang, Yun Wang, and Xiaoyu Chu are with the Department of Electronic Engineering, Shanghai Jiao Tong University, China. Jinbei Zhang is also with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, China. Xinbing Wang is also with National Mobile Communications Research Laboratory, Southeast University, Nanjing, China. E-mail: {abelchina, xwang8, xtian, sunshinehk, cxygrace}@sjtu.edu.cn.*
• *Yu Cheng is with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, USA. E-mail: {cheng}@iit.edu. Correspondence author: Dr. Xinbing Wang.*

the number of destinations per source. In their work, the network is partitioned into $\Theta(n)$ cells similar to [3] and TDMA scheme is used to avoid interference. Zhou and Ying [15] also studied the *fast mobility* model and provided an optimal tradeoff under their network assumptions. Specifically, they considered a network that consists of n_s multicast sessions, each of which had one source and p destinations. They showed that given delay constraint D, the capacity per multicast session was $O\left(\min\left\{1, (\log p)(\log(n_s p))\sqrt{\frac{D}{n_s}}\right\}\right)$. Then a *joint coding/scheduling algorithm* was proposed to achieve a throughput of $O\left(\min\left\{1, \sqrt{\frac{D}{n_s}}\right\}\right)$. In their network, each multicast session had no intersection with others and the total number of mobile nodes was $n = n_s(p+1)$.

Heterogeneous networks with multicast traffic pattern were studied by Li et al. [16] and Mao et al. [17]. Wired base stations are used and their transmission range can cover the whole network. Li et al. [18] studied a dense network with fixed unit area. The helping nodes in their work are wireless, but have higher power and only act as relays instead of sources or destinations. [16], [17] and [18] all study static networks.

In this paper, we give a general analysis on the optimal multicast capacity-delay tradeoffs in both homogeneous and heterogeneous MANETs. We assume a mobile wireless network that consists of n nodes, among which $n_s = n^s$ nodes are selected as sources and $n_d = n^\alpha$ destined nodes are chosen for each. Thus, n_s multicast sessions are formed. Our results in homogeneous network are further used to study the heterogeneous network, where $m = n^\beta$ base stations connected with wires are uniformly distributed in the unit square.

The purpose of this paper is to conduct extensive analysis on the multicast capacity-delay tradeoff in mobile wireless networks. We study a variety of mobility models

which are also widely adopted in previous works. The results obtained may provide valuable insights on how multicast will affect the network performance compared to unicast networks, e.g., [4] [5]. By removing some limitations and constraints, we try to present a fundamental and more general result than previous works.

We summarize our main results as follows where the logarithmic factors are omitted here:

1) Two-dimensional i.i.d. mobility models: Under a delay constraint D, the maximum throughput per multicast session is $O\left(\frac{n}{n_s n_d}\sqrt{\frac{D}{n}n_d}\right)$ for fast mobility, and $O\left(\frac{n}{n_s n_d}\sqrt[3]{\frac{D}{n}n_d}\right)$ for slow mobility.

2) Two-dimensional hybrid random walk mobility models: When $B = o(1)$ and $D = \omega(|\log B|/B^2)$, the maximum throughput per multicast session is $O\left(\frac{n}{n_s n_d}\sqrt{\frac{D}{n}n_d}\right)$ for fast mobility, and $O\left(\frac{n}{n_s n_d}\sqrt[3]{\frac{D}{n}n_d}\right)$ for slow mobility.

3) One-dimensional i.i.d. mobility models: Under a delay constraint D, the maximum throughput per multicast session is $O\left(\frac{n}{n_s n_d}\sqrt[3]{\frac{D^2}{n}n_d^2}\right)$ for fast mobility, and $O\left(\frac{n}{n_s n_d}\sqrt[4]{\frac{D^2}{n}n_d^2}\right)$ for slow mobility.

4) One-dimensional hybrid random walk mobility models: When $B = o(1)$ and $D = \omega(1/B^2)$, the maximum throughput per multicast session is $O\left(\frac{n}{n_s n_d}\sqrt[3]{\frac{D^2}{n}n_d^2}\right)$ for fast mobility, and $O\left(\frac{n}{n_s n_d}\sqrt[4]{\frac{D^2}{n}n_d^2}\right)$ for slow mobility.

5) Heterogeneous networks with infrastructure support:

   a) In infrastructure mode, the maximum aggregate input throughput of the whole network is $T_i = \min\{T_{i1}, T_{i2}\}$. The detail can be found in Section 10.

   b) The aggregate input capacity of the heterogeneous networks under the above eight different kinds of mobility models is $T = \max\left\{T_i, n_s\lambda_a\right\}$, where λ_a is the per session capacity in each of the homogeneous networks presented above.

The rest of the paper is organized as follows. In Section 2, we outline the system models. The eight mobility models in homogeneous networks are investigated in Section 3 to Section 8 respectively. Section 9 offers some discussions on the obtained results. In Section 10 we study the capacity of heterogeneous networks. In the end, we conclude this paper.

2 SYSTEM MODELS

We consider a mobile ad hoc network where n nodes move within a unit square. Among them, n_s nodes are selected as sources, and each node has n_d distinct destinations. We group each source and its n_d destinations as a multicast session. Note that a particular node may serves as both a source and a destination in different multicast sessions. Protocol Model [1] is employed. The definitions of capacity and delay are also similar to previous works, such as [1], [4] and [5].

2.1 Homogeneous Networks

Mobile ad hoc network model: Consider an ad hoc network where n wireless mobile nodes are randomly distributed in a unit square. The unit square is assumed to be a torus to avoid the border effect. We will study the following mobility models, similar to [5], in this paper.

1) **Two-dimensional i.i.d. mobility model:**

   a) At the beginning of each time slot, nodes will be uniformly and randomly distributed in the unit square.

   b) The node positions are independent of each other, and independent from time slot to time slot.

2) **Two-dimensional hybrid random walk model:** Consider a unit square which is further divided into $1/B^2$ squares of equal size. Each of the smaller square is called a RW-cell (random walk cell), and indexed by (U_x, U_y) where $U_x, U_y \in \{1, \ldots, 1/B\}$. A node which is in one RW-cell at a time slot moves to one of its eight adjacent RW-cells or stays in the same RW-cell in the next time-slot with a same probability. Two RW-cells are said to be adjacent if they share a common point. The node position within the RW-cell is randomly and uniformly selected.

3) **One-dimensional i.i.d. mobility model:**

   a) Reasonably, we assume the number of mobile nodes n and source nodes n_s are both even numbers. Among the mobile nodes, $n/2$ nodes (including $n_s/2$ source nodes), named H-nodes, move horizontally; and the other $n/2$ nodes (including the other $n_s/2$ source nodes), named V-nodes, move vertically.

   b) Let (x_i, y_i) denote the position of node i. If node i is a H-node, y_i is fixed and x_i is randomly and uniformly chosen from $[0, 1]$. We also assume that H-nodes are evenly distributed vertically, so y_i takes values $2/n, 4/n, \ldots, 1$. V-nodes have similar properties.

   c) Assume that source and destinations in the same multicast session are the same type of nodes. Also assume that node i is a H-node if i is odd, and a V-node if i is even.

   d) The orbit distance of two H(V)-nodes is defined to be the vertical (horizontal) distance of the two nodes.

4) **One-dimensional hybrid random walk model:** Each orbit is divided into $1/B$ RW-intervals (random walk interval). At each time slot, a node moves into one of two adjacent RW-intervals or stays at the current RW-interval. The node position in the RW-interval is randomly, uniformly selected.

We further assume that at each time slot, at most W bits can be transmitted in a successful transmission.

Mobility time scales: Two time scales of mobility are considered in this paper:

- **Fast mobility**: The mobility of nodes is at the same time scale as the transmission of packets, i.e., in each time-slot, only one transmission is allowed.
- **Slow mobility**: The mobility of nodes is much slower than the transmission of packets, i.e., multiple transmissions may happen within one time-slot.

Scheduling Policies: We assume that there exists a scheduler that has all the information about the current and past status of the network, and can schedule any radio transmission in the current and future time slots, similar to [4]. We say a packet p is successfully delivered if and only if all destinations within the multicast session have received the packet. In each time slot, for each packet p that has not been successfully delivered and each of its unreached destination k, the scheduler needs to perform the following two functions:

- **Capture**: The scheduler needs to decide whether to deliver packet p to destination k in the current time slot. If yes, the scheduler then needs to choose one relay node (possibly the source node itself) that has a copy of the packet p at the beginning of the time-slot, and schedules radio transmissions to forward this packet to destination k within the same time-slot, using possibly multi-hop transmissions. When this happens successfully, we say that the chosen relay node has successfully *captured* the destination k of packet p. We call this chosen relay node the *last mobile relay* for packet p and destination k. And we call the distance between the last mobile relay and the destination as the *capture range*.
- **Duplication**: For a packet p that has not been successfully delivered, the scheduler needs to decide whether to *duplicate* packet p to other nodes that do not have the packet at the beginning of the time-slot. The scheduler also needs to decide which nodes to relay from and relay to, and how.

2.2 Heterogeneous Networks

We introduce m regularly placed base stations (connected with each other via wires) into the mobile ad hoc networks and generate a heterogeneous network. Specifically, the base stations are placed at positions $(\frac{1}{2\sqrt{m}}+i\frac{1}{\sqrt{m}}, \frac{1}{2\sqrt{m}}+j\frac{1}{\sqrt{m}})$ with $0 \le i, j \le \sqrt{m}-1$. Clearly, these m regularly distributed base stations divide the original square region into m subregions with side length $\frac{1}{\sqrt{m}}$. Here we assume that m is the square of some integer for simplicity.

All transmissions can be carried out either in *ad hoc mode* or in *infrastructure mode* (see Figure 1). We assume that the base stations have a same transmission bandwidth, denoted by W_i for each. The bandwidth for each mobile ad hoc node is denoted by W_a. Further, we evenly divide the bandwidth W_i into two parts, one for uplink

transmissions and the other for downlink transmissions, so that these different kinds of transmissions will not interfere with each other.

A transmission in infrastructure mode is carried out in the following steps:

1) *Uplink:* A mobile node holding packet p is selected, and transmits this packet to the nearest base station.
2) *Infrastructure relay:* Once a base station receives a packet from a mobile node, all the other $m-1$ base stations share this packet immediately, (i.e., the delay is considered to be zero) since all base stations are connected by wires.
3) *Downlink:* Each base station searches for all the packets needed in its own subregion, and transmit all of them to their destined mobile nodes. At this step, every base station will adopt TDMA schemes to delivere different packets for different multicast sessions.
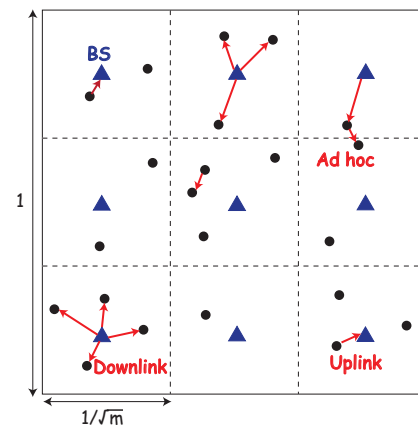

Fig. 1. Heterogeneous network with infrastructure support.

3 TWO DIMENSIONAL I.I.D. FAST MOBILITY MODEL

In this section, we present the upper bound on multicast capacity-delay tradeoff under the *two-dimensional i.i.d. fast mobility model*, and then propose a scheme to achieve a capacity close to the upper bound up to a logarithmic factor.

3.1 Upper Bound

Consider packet p and one of its destinations k, let $L_{p,k}$ denote the *capture range* for packet p and destination k, L_p denote the *capture range* for packet p and its last reached destination. Let $D_{p,k}$ denote the number of time slots it takes to reach destination k after reaching destination $k-1$. Denote D_p as the number of time slots it takes to reach the last destination of packet p, which means $D_p = \sum_k D_{p,k}$. And let $R_{p,k}$ denote the number of mobile relays holding packet p when the packet reaches

its k-th destination and last destination respectively. To reach a new destination k, all the nodes holding packet p should move across a fraction of network area in $D_{p,k}$ timeslots. Then we have the following lemma.

Lemma 1: Under *two-dimensional i.i.d. mobility model* and concerning successful encounter, the following inequality holds for any causal scheduling policy,

$$c_1 \log n \mathbb{E}[D_{p,k}] \geq \frac{1}{(n_d - k + 1)\left(\mathbb{E}[L_{p,k}] + \frac{1}{n^2}\right)^2 \mathbb{E}[R_{p,k}]} \tag{1}$$

where c_1 is a positive constant.

Consider a large time interval T. The total number of packets communicated among all sessions is $\lambda n_s T$. Then we have the following lemma.

Lemma 2: Under *fast mobility model* and concerning network radio resource consumption, the following inequality holds for any causal scheduling policy (c_2 is a positive constant),

$$\sum_{p=1}^{\lambda n_s T} \frac{\Delta^2}{4} \frac{\mathbb{E}[R_p] - n_d}{n} + \sum_{p=1}^{\lambda n_s T} \sum_{k=1}^{n_d} \frac{\pi \Delta^2}{4} \mathbb{E}[L_{p,k}^2] \leq c_2 W T \log n. \tag{2}$$

Proof: Here are some intuitive explanations. Proofs are similar to and can be easily inferred from Appendix B in [4].

We try to measure how much radio resource each transmission consumes, by calculating the areas of disjoint disks caused by interference. Radio resource consumption is divided into two parts, *Capture* and *Duplication*.

- *Capture*: For each packet p and each of its destination k, the *one-hop capture*[1] consumes an area of $\frac{\pi \Delta^2}{4}(L_{p,k})^2$. Hence, the upper bound on the expected area consumed by all n_d successful captures of packet p is $\sum_{k=1}^{n_d} \frac{\pi \Delta^2}{4} \mathbb{E}[L_{p,k}^2]$.
- *Duplication*: If the radius of transmission range is s, then w.h.p. there are $\pi s^2 n$ nodes which can receive the broadcast packets, and a disk of area $\frac{\pi \Delta^2}{4} s^2$ centered at the transmitter will be disjoint from others. Therefore, we can use $\frac{\Delta^2}{4} \frac{\mathbb{E}[R_p] - n_d}{n}$ as an upper bound on the expected area consumed by producing $R_p - n_d$ copies of the packet to other nodes before any of them or the source itself successfully forwards the packet to the last destination. Note that since we use cooperative mode [14], where destinations can also act as relays, the copies produced in *Duplication* should not only exclude the source node but also exclude the $n_d - 1$ destinations which receive the copies in *Capture* procedure.

□

Theorem 1: Under *two-dimensional i.i.d. fast mobility model*, let D be the mean delay averaged over all packets, and let λ be the capacity per multicast session. The

following upper bound holds for any causal scheduling policy,

$$\lambda \leq \min \left\{ \Theta(1), \Theta\left(\frac{n}{n_s n_d} \sqrt{\frac{n_d D}{n} \log^3 n}\right) \right\} \tag{3}$$

Proof: Each source can send out at most W size of packet per time-slot, i.e., $\lambda \leq W = \Theta(1)$. Therefore, we only need to prove the second part.

From *Lemma 1*, we have

$$\begin{aligned}
\sum_{k=1}^{n_d}(\mathbb{E}[L_{p,k}] + \frac{1}{n^2})^2 &\geq \sum_{k=1}^{n_d} \frac{1}{(n_d - k + 1)\mathbb{E}[R_{p,k}]\mathbb{E}[D_{p,k}] \log n} \\
&\geq \frac{1}{\mathbb{E}[R_p] \log n} \sum_{k=1}^{n_d} \frac{1}{(n_d - k + 1)\mathbb{E}[D_{p,k}]} \\
&\geq \frac{1}{\mathbb{E}[R_p] \log n} \frac{\left(\sum_{k=1}^{n_d} \frac{1}{\sqrt{n_d-k+1}}\right)^2}{\sum_{k=1}^{n_d} \mathbb{E}[D_{p,k}]} \\
&\geq \frac{1}{\mathbb{E}[R_p] \log n} \cdot \frac{c_2' n_d}{\mathbb{E}[D_p]},
\end{aligned} \tag{4}$$

where c_2' is a constant. Note that $\left(\sum_{k=1}^{n_d} \frac{1}{\sqrt{n_d-k+1}}\right)^2 = \Theta(n_d)$ when $n_d = \Omega(1)$. Equations will hold when $R_{p,k} = \Theta(R_p)$ and $\sqrt{n_d - k + 1} D_{p,k} = \Theta(\sqrt{n_d - i + 1} D_{p,i})$ for all k and i.

There are two cases we need to consider.
Case 1: When $\sum_{k=1}^{n_d} \mathbb{E}[L_{p,k}^2] = \Omega(\frac{1}{n^4})$, from *Lemma 2*,

$$\begin{aligned}
&\sum_{p=1}^{\lambda n_s T} \frac{\Delta^2}{4} \frac{\mathbb{E}[R_p]}{n} + \sum_{p=1}^{\lambda n_s T} \sum_{k=1}^{n_d} \frac{\pi \Delta^2}{4} \mathbb{E}[L_{p,k}^2] \\
&\geq \sum_{p=1}^{\lambda n_s T} \frac{\Delta^2}{4} \frac{\mathbb{E}[R_p]}{n} + c_3' \sum_{p=1}^{\lambda n_s T} \frac{1}{\mathbb{E}[R_p] \log n} \cdot \frac{n_d}{\mathbb{E}[D_p]} \\
&\geq c_1' \sum_{p=1}^{\lambda n_s T} \sqrt{\frac{n_d}{n \mathbb{E}[D_p] \log n}} \\
&= c_1' \sqrt{\frac{n_d}{n \log n}} \sum_{p=1}^{\lambda n_s T} \sqrt{\frac{1}{\mathbb{E}[D_p]}} \\
&\geq c_1' \sqrt{\frac{n_d}{n \log n}} \frac{(\sum_{p=1}^{\lambda n_s T} 1)^2}{\sum_{p=1}^{\lambda n_s T} \sqrt{\mathbb{E}[D_p]}} \\
&\geq c_1' \sqrt{\frac{n_d}{n \log n}} \frac{(\sum_{p=1}^{\lambda n_s T} 1)^2}{\sqrt{\sum_{p=1}^{\lambda n_s T} \mathbb{E}[D_p] \cdot \sum_{p=1}^{\lambda n_s T} 1}} \\
&= c_1' \sqrt{\frac{n_d}{n \log n}} \frac{\lambda n_s T}{\sqrt{D}}
\end{aligned} \tag{5}$$

where c_1' and c_3' are both constants. Equations hold when $R_p = \Theta(\sqrt{\frac{n n_d}{D_p \log n}})$ and $D_p = \Theta(D_i)$. Therefore,

$$\sqrt{\frac{n_d}{n \log n}} \frac{\lambda n_s T}{\sqrt{D}} - \frac{\lambda n_s T n_d}{n} \leq W T \log n. \tag{6}$$

---

1. Concerning the *multi-hop capture*, consumption area is summed up by each hop transmission.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON MOBILE COMPUTING

5

When $D = O(\frac{n}{n_d \log n})$, the first term dominates and hence,

$$\lambda \leq \frac{\sqrt{nD \log^3 n}}{n_s \sqrt{n_d}}. \tag{7}$$

Case 2: When $\sum_{k=1}^{n_d} \mathbb{E}[L_{p,k}^2] = O(\frac{1}{n^4})$, the first term in the left part of *Lemma 2* would dominate. We have

$$\frac{4c_2 WT \log n}{\Delta^2} \geq \frac{1}{4c_1 \log n} \frac{\lambda T n_s n^3}{D} - \lambda T \frac{n_s n_d}{n}.$$

Hence, for n large enough,

$$\lambda \leq \frac{32 c_1 c_2 W}{\Delta^2} \frac{D \log^2 n}{n_s n^3}. \tag{8}$$

Finally, we compare the two inequalities we have obtained, i.e., (4) and (8). When $D = O(\frac{n}{n_d \log n})$, inequality (8) will eventually be the loosest for large n, the optimal capacity-delay tradeoff is upper bounded by

$$\lambda \leq \Theta\left(\frac{n}{n_s n_d} \sqrt{\frac{D \log^3 n}{n} n_d} \right).$$

$\square$

3.2 Achievable Lower Bound

In this subsection, we will show how the study of the upper bound also helps us in developing a new scheme that can achieve a capacity-delay tradeoff that is close to the upper bound.

Choosing Optimal Values of Key Parameters:
From *Theorem 1*, we have

$$\lambda = O\left(\frac{n}{n_s n_d} \sqrt{\frac{n_d D \log^3 n}{n}} \right) = O(n^{-\frac{2s+\alpha-1-d}{2}} \log^{\frac{3}{2}} n).$$

In order to achieve the maximum capacity on the right hand side, all inequalities in the proof of Theorem 1 should hold with equality. By studying the conditions under which these inequalities are tight, we are able to identify that the optimal choices of various key parameters of the scheduling policy. We can infer that the parameters (such as $\mathbb{E}[R_{p,k}]$, $\mathbb{E}[L_{p,k}]$) of each packet p and each destination node k should be the same and concentrate on their respective average values. This implies that the scheduling policy should use the same parameters for all packets and all destinations. We further assume that $n_s = n^s, 0 \leq s \leq 1$; $n_d = n^\alpha, 0 \leq \alpha \leq 1$ and $D = n^d, 0 \leq d < 1 - \alpha$. In addition, we assume the number of mobile nodes $n \leq n_s n_d$. This notation is used throughout all other tables in this paper. The results are summarized in Table 1.

Capacity Achieving Scheme I:
We propose a flexible cell-partitioning scheme to achieve a capacity that is close to the upper bound, using broadcasting and time division. Cell-partitioning schemes, like [3] and [4], divide the network into several non-overlapping and independent cells and only allow transmissions within the same cell. As *Lemma 2* in [18]

TABLE 1
The order of the optimal values of the parameters in two-dimensional fast i.i.d. mobility model.

L: Capture Range	$\Theta(n^{-\frac{1+\alpha+d}{4}} / \log^{\frac{1}{4}} n)$
R: # of Duplicates	$\Theta(n^{\frac{1+\alpha-d}{2}} / \log^{\frac{1}{2}} n)$

shows, each cell in the network can transmit at a rate of $c_3 W$, where c_3 is a deterministic positive constant.

We group every D time-slots into a super-slot.

1) At each odd super-slot, we schedule transmissions from the sources to the relays in every time-slot. We divide the unit square into $\mathcal{C}_d = \Theta\left(\frac{n^{(1-\alpha+d)/2}}{\log n} \right)$ cells. Each cell is a square of area $1/\mathcal{C}_d$. We refer to each cell in the odd super-slot as a *duplication cell*. By *Lemma 6* in [4], each cell can be active for $1/c_4$ amount of time, where c_4 is some constant. When a cell is scheduled to be active, each source node in the cell broadcasts a new packet to all other nodes in the same cell for $\Theta\left(\frac{n^{-(2s+\alpha-1-d)/2}}{\log^2 n} \right)$ amount of time. These other nodes then serve as mobile relays for the packet. The nodes within the same *duplication cell* coordinate themselves to broadcast sequentially.

2) At each even super-slot, we schedule transmissions from the mobile relays to the destination nodes in every time-slot. We divide the unit square into $\mathcal{C}_c = \Theta\left(n^{(1+\alpha+d)/2} \right)$ cells. Each cell is a square of area $1/\mathcal{C}_c$. We refer to each cell in the even super-slot as the *capture cell*. In each time-slot, for each destination node $\mathcal{D}$ and each of its source node $\mathcal{S}$, pick a node $Y_{\mathcal{S}\mathcal{D}}$ that is in the same *capture cell* with node $\mathcal{D}$ in current time-slot and in the same *duplication cell* with node $\mathcal{S}$ some time-slot in previous super-slot and hold a copy of the packet source node $\mathcal{S}$. If there are multiple relay nodes, just pick one, which we call a *representative relay*, and transmit the destined packet to $\mathcal{D}$. At the end of each even super-slot, clear all the buffers of mobile nodes, and prepare for a new turn of duplication and capture.

As $n \to \infty$, with high probability (w.h.p.), all packets generated in odd duplication super-slot will finish its n_d destined transmission within the following even capture super-slot.

Proposition 1: With probability approaching one, as $n \to \infty$, the above scheme allows each source to send D packets of size $\lambda = \Theta\left(\frac{n^{-(2s+\alpha-1-d)/2}}{\log^2 n} \right)$ to their respective destinations within $2D$ time-slots.

4 TWO DIMENSIONAL I.I.D. SLOW MOBILITY MODEL

In this section, we present the upper bound on multicast capacity-delay tradeoff under the *two-dimensional i.i.d. slow mobility model*, and then propose a scheme to achieve

a capacity close to the upper bound up to logarithmic factors.

4.1 Upper Bound

Under slow mobility model, once a successful capture with respect to packet p and one of its destination k occurs, the last mobile relay will start transmitting packet p to destination k within a single time slot, using possibly other nodes as relays. Let $h_{p,k}$ denote the number of hops packet p taken from the last mobile relay to destination k. And let $S_{p,k}^h$, $h = 1, 2, \ldots, h_{p,k}$ denote the length of each hop. Hence, similar to *Lemma 2*, the following lemma holds.

Lemma 3: Under *slow mobility model* and concerning network radio resources consumption, the following inequality holds for any causal scheduling policy (c_4 is some positive constant).

$$\sum_{p=1}^{\lambda n_s T} \frac{\Delta^2}{4} \frac{\mathbb{E}[R_p] - n_d}{n} + \sum_{p=1}^{\lambda n_s T} \sum_{k=1}^{n_d} \sum_{h=1}^{h_{p,k}} \frac{\pi \Delta^2}{4} \mathbb{E}\left[(S_{p,k}^h)^2\right]$$
$$\leq c_5 WT \log n, \quad (9)$$

where the sum of the hop's lengths of the $h_{p,i}$ hops must be no smaller than the straight-line distance-capture radius:

$$\sum_{h=1}^{h_{p,k}} S_{p,k}^h \geq L_{p,k}. \quad (10)$$

Theorem 2: Under *two-dimensional i.i.d. slow mobility model*, let D be the mean delay averaged over all packets, and let λ be the capacity per multicast session. The following upper bound holds for any causal scheduling policy,

$$\lambda = O\left(\frac{n \log n}{n_s n_d} \sqrt[3]{\frac{n_d D}{n}}\right) \quad (11)$$

Proof: Since some of the arguments are similar to previous sections, we only present the main proof here.

A node can either transmit or receive at one time. Therefore, it is easy to see that,

$$\sum_{p=1}^{\lambda n_s T} \sum_{k=1}^{n_d} \sum_{h=1}^{h_{p,k}} 1 \leq \frac{WT}{2} n. \quad (12)$$

Using Cauchy-Schwartz inequality and (12), we have

$$\sum_{p=1}^{\lambda n_s T} \sum_{k=1}^{n_d} \sum_{h=1}^{h_{p,k}} \mathbb{E}\left[(S_{p,k}^h)^2\right] = \mathbb{E}\left[\sum_{p=1}^{\lambda n_s T} \sum_{k=1}^{n_d} \sum_{h=1}^{h_{p,k}} (S_{p,k}^h)^2\right]$$
$$\geq \frac{2}{WTn}\left(\mathbb{E}\left[\sum_{p=1}^{\lambda n_s T} \sum_{k=1}^{n_d} \sum_{h=1}^{h_{p,k}} S_{p,k}^h\right]\right)^2$$
$$\geq \frac{2}{WTn}\left(\mathbb{E}\left[\sum_{p=1}^{\lambda n_s T} \sum_{k=1}^{n_d} L_{p,k}\right]\right)^2 \quad (13)$$

Equalities hold when (12) becomes an equality and $S_{p,k}^h$ is equal for all p, k and h.

From *Lemma 1*, we have

$$\sum_{k=1}^{n_d} L_{p,k} \geq \sum_{k=1}^{n_d} \frac{1}{\sqrt{R_p \log n D_{p,k}(n_d - k)}}$$
$$= \frac{1}{\sqrt{R_p \log n}} \sum_{k=1}^{n_d} \frac{1}{\sqrt{D_{p,k}(n_d - k)}}$$
$$\geq \frac{1}{\sqrt{R_p \log n}} \frac{\sum_{k=1}^{n_d}(n_d - k)^{-\frac{1}{4}}}{\sum_{k=1}^{n_d} \sqrt{D_{p,k}}} \quad (14)$$
$$\geq \frac{1}{\sqrt{R_p \log n}} \frac{n_d^{\frac{3}{2}}}{\sqrt{\sum_{k=1}^{n_d} D_{p,k} \sum_{k=1}^{n_d} 1}}$$
$$= \frac{1}{\sqrt{R_p \log n}} \frac{n_d}{\sqrt{D_p}}.$$

Note that,

$$\sum_{p=1}^{\lambda n_s T} \sum_{k=1}^{n_d} L_{p,k} \geq \sum_{p=1}^{\lambda n_s T} \frac{1}{\sqrt{R_p \log n}} \frac{n_d}{\sqrt{D_p}}$$
$$\geq \frac{n_d}{\sqrt{\log n}} \frac{(\sum_{p=1}^{\lambda n_s T} R_p^{-\frac{1}{4}})^2}{\sum_{p=1}^{\lambda n_s T} \sqrt{D_p}}$$
$$\geq \frac{n_d}{\sqrt{\log n}} \frac{(\sum_{p=1}^{\lambda n_s T} R_p^{-\frac{1}{4}})^2}{\sqrt{\sum_{p=1}^{\lambda n_s T} D_p \sum_{p=1}^{\lambda n_s T} 1}} \quad (15)$$
$$= c_3 \frac{(\sum_{p=1}^{\lambda n_s T} R_p^{-\frac{1}{4}})^2}{\sqrt{D}}$$

where $c_3 = \frac{n_d}{\sqrt{\log n \lambda n_s T}}$.

$$\sum_{p=1}^{\lambda n_s T} \frac{\mathbb{E}[R_p]}{n} + \sum_{p=1}^{\lambda n_s T} \sum_{k=1}^{n_d} \sum_{h=1}^{h_{p,k}} \mathbb{E}\left[(S_{p,k}^h)^2\right]$$
$$\geq \sum_{p=1}^{\lambda n_s T} \frac{R_p}{n} + \frac{2}{WTn} \frac{c_3^2}{D}\left(\sum_{p=1}^{\lambda n_s T} R_p^{-\frac{1}{4}}\right)^4. \quad (16)$$

Applying Cauchy-Schwartz inequality again and again, it can be proved that Equation (16) is greater than $\Theta\left(\sqrt{\frac{\lambda^3 n_s^3 T^2 n_d^2}{\log n D n^2}}\right)$.

Hence, the optimal capacity-delay tradeoff is upper bounded by

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

$\square$

4.2 Achievable Lower Bound

Choosing Optimal Values of Key Parameters:
From *Theorem 2*, we have

$$\lambda = O\left(\frac{n}{n_s n_d} \sqrt[3]{\frac{n_d D \log^3 n}{n}}\right) = O(n^{-\frac{3s+2\alpha-2-d}{3}} \log n)$$

The idea is similar, as is presented in Section 3.2. We summarize the *optimal values* in Table 2.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON MOBILE COMPUTING

7

TABLE 2
The order of the optimal values of the parameters in two-dimensional slow i.i.d. mobility model.

L: Capture Range	$\Theta(n^{-\frac{1+2\alpha+2d}{6}}/\log^{\frac{1}{2}} n)$
R: # of Duplicates	$\Theta(n^{\frac{1+2\alpha-d}{3}})$
H: # of Hops	$\Theta(n^{\frac{1-\alpha-d}{3}}/\log n)$
S: Hop Length	$\Theta(\sqrt{\log n/n})$

Capacity Achieving Scheme II: We group every D time-slots into a super-slot. Scheme II is similar to Capacity Achieving Scheme I presented in Section 3.2, and we only introduce the differences here.

1) At each odd super-slot, we schedule transmissions from the sources to the relays in every time-slot. We divide the unit square into $\mathcal{C}_d = \Theta\left(\frac{n^{(2-2\alpha+d)/3}}{\log n}\right)$ cells. When a cell is scheduled to be active, each source node in the cell broadcasts a new packet to all other nodes in the same cell for $\Theta\left(\frac{n^{-(3s+2\alpha-2-d)/3}}{\log^2 n}\right)$ amount of time.

2) At each even super-slot, we schedule transmissions from the mobile relays to the destination nodes in every time-slot. We divide the unit square into $\mathcal{C}_c = \Theta\left(n^{(1+2\alpha+2d)/3}\right)$ cells. After picking out a representative relay, we then schedule multi-hop transmissions in the following fashion to forward each packet from the *representative relay* to its destination in the same *capture cell*. We further divide each *capture cell* into $\mathcal{C}_h = \Theta\left(\frac{n^{(2-2\alpha-2d)/3}}{\log n}\right)$ hop-cells (in $\sqrt{\mathcal{C}_h}$ rows and $\sqrt{\mathcal{C}_h}$ columns). Each hop-cell is a square of area $1/(\mathcal{C}_c\mathcal{C}_h)$. By *Lemma 6* in [4], there exists a scheduling scheme where each hop-cell can be active for $1/c_4$ amount of time. When each hop-cell is active, it forwards a packet to another node in the neighboring hop-cell. If the destination of the packet is in the neighboring cell, the packet is forwarded directly to the destination node. The packets from each *representative relay* are first forwarded towards neighboring cells along the X-axis, then to their destination nodes along the Y-axis. At the end of each even super-slot, clear all the buffers of mobile nodes, and prepare for a new turn of duplication and capture.

Proposition 2: With probability approaching one, as $n \to \infty$, the above scheme allows each source to send D packets of size $\lambda = \Theta\left(\frac{n^{-(3s+2\alpha-2-d)/3}}{\log^2 n}\right)$ to their respective destinations within $2D$ time-slots.

Proof: We will focus on the case in which the mean delay is bounded by a constant, i.e., $D = 1$. Let $\lfloor x \rfloor$ be the largest integer smaller than or equal to x. We use the following values[2]: $\mathcal{C}_d = \lfloor(\frac{n^{(2-2\alpha)/3}}{8\log n})^{\frac{1}{2}}\rfloor^2$, $\mathcal{C}_c = \lfloor(n^{(1+2\alpha)/3})^{\frac{1}{2}}\rfloor^2$, $\mathcal{C}_h = \lfloor(\frac{n^{(2-2\alpha)/3}}{4\log n})^{\frac{1}{2}}\rfloor^2$. We will show that

2. To ensure the positive values, we assume $\alpha < 1$.

our scheme can obtain a capacity of $\frac{W}{32n^{(3s+2\alpha-2)/3}\log n}$ w.h.p. under *multicast* traffic pattern. First, we present a lemma which will be used frequently in later proof. It has already been proven in Lemma 11, [4].

Lemma 4: Consider an experiment where we randomly throw n balls into $m \leq n$ independent urns. The success probability for each ball to enter any one of the urns is $p \leq 1$. Let $B_i, i = 1, \ldots, m$ be the number of balls in urn i after n balls are thrown. Then $\mathbb{E}[B_i] = \frac{np}{m}$. And as $n \to \infty$, we have

1) If $\frac{np}{m} \geq c \log n$, and $c \geq 8$, then

$$\mathbf{P}[B_i \geq 2\frac{np}{m} \text{ for any } i] \leq \frac{1}{n}$$

2) If $\frac{np}{m} \geq cn^{\alpha}$, where $c > 0$ and $\alpha > 0$, then

$$\mathbf{P}[B_i \geq 2\frac{np}{m} \text{ for any } i] = O(\frac{1}{n})$$

3) If $\frac{np}{m} \geq c \log n$ and $c \geq 4$, then

$$\mathbf{P}[B_i = 0 \text{ for any } i] = O(\frac{1}{n})$$

Analysis of *Duplication*

We consider the experiment in which we throw n_s balls into $\mathcal{C}_d$ urns with $p = 1$. We have

$$16n^{(3s+2\alpha-2)/3}\log n \geq \frac{n_s}{\mathcal{C}_d} \geq 8n^{(3s+2\alpha-2)/3}\log n.$$

Let $N_d(i)$ denote the number of source nodes in *duplication cell* i. Since $n \leq n_s n_d$, i.e., $s + \alpha \geq 1$, by *Lemma 4* (1), we have

$$\mathbf{P}[N_d(i) \geq 32n^{(3s+2\alpha-2)/3}\log n \text{ for any } i]$$
$$\leq \quad \mathbf{P}[N_d(i) \geq 2\frac{n_s}{\mathcal{C}_d} \text{for any } i] \leq \frac{1}{n}.$$

Hence, w.h.p., there are no more than $32n^{(3s+2\alpha-2)/3}\log n$ source nodes within the same *duplication cell*. Using time division, we can arrange each source to broadcast a packet for $\frac{1}{32n^{(3s+2\alpha-2)/3}\log n}$ amount of time in sequence.

Analysis of *Capture*

We consider the experiment in which we throw n balls into $\mathcal{C}_c\mathcal{C}_h$ urns with $p = 1$. We have

$$\frac{n}{\mathcal{C}_c\mathcal{C}_h} \geq 4\log n.$$

Let $N_h(i)$ denote the number of nodes in *hopping cell* i. By *Lemma 4* (3), we have

$$\mathbf{P}[N_h(i) = 0 \text{ for any } i] = O(\frac{1}{n}).$$

Hence, w.h.p., there is always a node in each *hopping cell* that helps the multi-hop transmission.

Then we consider the experiment in which we throw n balls into $\mathcal{C}_d\mathcal{C}_c$ urns with $p = 1$. We have

$$16\log n \geq \frac{n}{\mathcal{C}_d\mathcal{C}_c} \geq 8\log n.$$

Let $N_{dc}(i, j)$ denote the number of nodes that are in *duplication cell* i in the previous time-slot and now in

capture cell j in the current time-slot. By *Lemma 4* (3), we have

$$\mathbf{P}[N_{dc}(i,j) = 0 \text{ for any } i,j] = O(\frac{1}{n}).$$

Hence, w.h.p., in each *capture cell* j, there is always a node which used to be in *duplication cell* i, and it has all the packets broadcast in that duplication cell. If there are multiple satisfying nodes, we only pick one for each i as the *representative relay* and we get $\mathcal{C}_d$ *representative relays* in *capture cell* j. On the other hand, all packets can be found in each *capture cell*, and each destination node can find all the destined packets it desires from the *representative relays*. Hence, we only need to calculate the maximum transmissions passing through each *hopping cell* if all desired transmissions are allowed.

We consider the possible *transmission pairs* instead of the actual mobile nodes. A transmission pair is defined as the transmitting node and receiving node in a transmission. We classify the destinations based on the sessions they belong to, i.e., one destination may be calculated multiple times when it belongs to different sessions. Thus, there are $n_s n_d$ number of pairs either with different transmission nodes or requiring different packets.

For the transmissions horizontally passing through the *hopping cell*, we consider the experiment in which we throw $n_s n_d$ balls into $\mathcal{C}_d \mathcal{C}_c$ urns with $p = 1$. We have

$$16n^{s+\alpha-1} \log n \geq \frac{n_s n_d}{\mathcal{C}_d \mathcal{C}_c} \geq 8n^{s+\alpha-1} \log n.$$

Let $N_s(i,j)$ denote the number of *transmission pairs* whose source nodes are located in *duplication cell* i and destination nodes are located in *capture cell* j. By *Lemma 4* (1), we have

$$\mathbf{P}[N_s(i,j) \geq 32n^{s+\alpha-1} \log n \text{ for any } i,j]$$
$$\leq \mathbf{P}[N_s(i,j) \geq 2\frac{n_s n_d}{\mathcal{C}_d \mathcal{C}_c} \text{ for any } i,j] \leq \frac{1}{n_s n_d}.$$

Hence, w.h.p., in each *capture cell* j, each *representative relay* will serve no more than $32n^{s+\alpha-1} \log n$ *transmission pairs*.

Since $\mathcal{C}_d$ *representative relays* are chosen in each *capture cell*, we consider the experiment where we throw $\mathcal{C}_d$ balls into $\sqrt{\mathcal{C}_h}$ urns with $p = 1$. We have

$$\frac{\sqrt{2}n^{(1-\alpha)/3}}{4\sqrt{\log n}} \geq \frac{\mathcal{C}_d}{\sqrt{\mathcal{C}_h}} \geq \frac{n^{(1-\alpha)/3}}{8\sqrt{\log n}}.$$

Let $N_r(l)$ denote the number of *representative relays* in row l. Since $\alpha < 1$, by *Lemma 4* (2), we have

$$\mathbf{P}[N_r(j,l) \geq \frac{\sqrt{2}n^{(1-\alpha)/3}}{2\sqrt{\log n}} \text{ for any } j,l]$$
$$\leq \mathbf{P}[N_r(j,l) \geq 2\frac{\mathcal{C}_d}{\sqrt{\mathcal{C}_h}} \text{ for any } j,l] = O(\frac{1}{\mathcal{C}_d}) \to 0.$$

Hence, w.h.p., the number of horizontal transmissions $T_x = N_s(i,j)N_r(j,l) \leq 16n^{(3s+2\alpha-2)/3}\sqrt{2\log n}.$

For the transmissions vertically passing through the *hopping cell*, we consider the experiment where we throw $n_s n_d$ balls into $\mathcal{C}_c\sqrt{\mathcal{C}_h}$ urns with $p = 1$. We have

$$4n^{(3s+2\alpha-2)/3}\sqrt{2\log n} \geq \frac{n_s n_d}{\mathcal{C}_c\sqrt{\mathcal{C}_h}} \geq 2n^{(3s+2\alpha-2)/3}\sqrt{2\log n}.$$

Let $N_{ch}(j,l)$ denote the number of *transmission pairs* whose destinations are located in *capture cell* j and column l. By *Lemma 4* (2), we have

$$\mathbf{P}[N_{ch}(j,l) \geq 8n^{(3s+2\alpha-2)/3}\sqrt{2\log n}\text{for any } j,l]$$
$$\leq \mathbf{P}[N_{ch}(j,l) \geq 2\frac{n_s n_d}{\mathcal{C}_c\sqrt{\mathcal{C}_h}} \text{ for any } l] = O(\frac{1}{n_s n_d}).$$

Hence, w.h.p., the number of vertical transmissions is $T_y = N_{ch}(j,l) \leq 8n^{(3s+2\alpha-2)/3}\sqrt{2\log n}$. And the total transmissions passing through a single *hopping cell* are $T_x + T_y \leq 32n^{(3s+2\alpha-2)/3} \log n.$ □

5 TWO DIMENSIONAL HYBRID R.W. MOBILITY MODEL

In this section, we study the two-dimensional hybrid random walk mobility model with both fast and slow mobiles. We will obtain the maximum throughput for $D = \omega(|\log B|/B^2)$.

From Appendix G in [5], we have the following lemma.

Lemma 5: Under *two-dimensional hybrid random walk mobility model*, when given delay constraint $D = \omega(|\log B|/B^2)$, for any $L \in [0, B/\sqrt{\pi})$, we have

$$Pr(\widetilde{L}_p \leq L) \leq 36L^2 D, \tag{17}$$

where $\widetilde{L}_p$ is the minimum distance between a particular mobile relay of packet p and one of its destinations within D time slots.

Compared with two-dimensional i.i.d. mobility model, where $Pr(\widetilde{L}_p \leq L) = 1-(1-\pi L^2)^D \leq \pi L^2 D$, Lemma 5 is different only in the coefficient, which does not influence the orders of the final result. So following the same proof procedure of Theorem 1 and Theorem 2, we have the following two results in two-dimensional hybrid random walk model with fast and slow mobiles respectively.

Theorem 3: Under *two-dimensional hybrid random walk fast mobility model*, let D be the mean delay averaged over all packets, and let λ be the capacity per multicast session. When $B = o(1)$ and $D = \omega(|\log B|/B^2)$, the following upper bound holds for any causal scheduling policy,

$$\lambda = O\left(\frac{n}{n_s n_d}\sqrt{\frac{n_d D}{n}\log^3 n}\right). \tag{18}$$

Theorem 4: Under *two-dimensional hybrid random walk slow mobility model*, let D be the mean delay averaged over all packets, and let λ be the capacity per multicast session. When $B = o(1)$ and $D = \omega(|\log B|/B^2)$, the following upper bound holds for any causal scheduling policy,

$$\lambda = O\left(\frac{n \log n}{n_s n_d}\sqrt[3]{\frac{n_d D}{n}}\right). \tag{19}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON MOBILE COMPUTING

9

As the two-dimensional random walk mobility model has the same capacity upper bound as two-dimensional i.i.d. mobility model, Capacity Achieving Scheme I and Scheme II still apply to R.W. mobility model with fast and slow mobiles respectively, with some extra limitations on delay constraint. We do not extend this part due to space limitations.

6 ONE DIMENSIONAL I.I.D. FAST MOBILITY MODEL

6.1 Upper Bound

Lemma 6: Under *one-dimensional i.i.d. mobility model* and concerning successful encounter, the following inequality holds for any causal scheduling policy (c_6 is some positive constant).

$$c_6 \log n \mathbb{E}[D_p] \geq \frac{1}{\left(\mathbb{E}[L_p] + \frac{1}{n}\right)\mathbb{E}[R_p]}. \tag{20}$$

Proof:

To proof this lemma, we will need the following lemma on the minimum distance between the mobile relays and the destination at any time slot. Fix a packet p that enters into the system at time slot $t_0(p)$, and one of its destinations k. At each time slot $t \geq t_0(p)$, let $r_p(t)$ denote the number of mobile relays holding the packet p at the beginning of the time slot t. Among these $r_p(t)$ mobile relays, there is one mobile relay whose distance to the destination k of packet p is the smallest. Let $\widetilde{L}_{p,k}(t)$ denote this minimum distance, and let

$$L_{p,k}(t) = \max\{\frac{1}{n}, \widetilde{L}_{p,k}(t)\}$$

It is easy to verify that

$$\widetilde{l}_{p,k}(t) \geq \widetilde{L}_{p,k}(t) \geq L_{p,k}(t) - \frac{1}{n},$$

where $\widetilde{l}_{p,k}(t)$ is the distance between a chosen relay node holding packet p and the destination k.

Lemma 7: Under the one-dimensional i.i.d. mobility model, if $n \geq 3$, then

$$\mathbb{E}\left[\frac{1}{L_{p,k}(t)r_p(t)}|\mathcal{F}_{t-1}{}^3\right] \leq 8 \log n \text{ for all } t \leq t_0(p).$$

Proof: Let $\mathbf{I}_A$ be the indicator function on the set A. By the definition of $L_{p,k}(t)$, we have,

$$\mathbb{E}\left[\frac{1}{L_{p,k}(t)}|\mathcal{F}_{t-1}\right] = \mathbb{E}\left[n\mathbf{I}_{\{\widetilde{L}_{p,k}(t)\leq\frac{1}{n}\}}|\mathcal{F}_{t-1}\right]$$
$$+\mathbb{E}\left[\frac{1}{\widetilde{L}_{p,k}(t)}\mathbf{I}_{\{\widetilde{L}_{p,k}(t)>\frac{1}{n}\}}|\mathcal{F}_{t-1}\right].$$

3. Let $\mathcal{F}_t$ be the σ-algebra that captures all information about the "history" up to time-slot t, including the nodes' positions and the packets they have.

Since the nodes move on a unit square, $\widetilde{L}_{p,k}(t) \leq \sqrt{2}$. Hence,

$$\mathbb{E}\left[\frac{1}{\widetilde{L}_{p,k}(t)}\mathbf{I}_{\{\widetilde{L}_{p,k}(t)>\frac{1}{n}\}}|\mathcal{F}_{t-1}\right]$$
$$= \frac{1}{\sqrt{2}} - n\mathbf{P}[\widetilde{L}_{p,k}(t) \leq \frac{1}{n}|\mathcal{F}_{t-1}]$$
$$+ \int_{\frac{1}{n}}^{\sqrt{2}} \frac{1}{u^2}\mathbf{P}[\widetilde{L}_{p,k}(t) \leq u|\mathcal{F}_{t-1}]du.$$

Hence,

$$\mathbb{E}\left[\frac{1}{L_{p,k}(t)}|\mathcal{F}_{t-1}\right] = \frac{1}{\sqrt{2}} + \int_{\frac{1}{n}}^{\sqrt{2}} \frac{1}{u^2}\mathbf{P}[\widetilde{L}_{p,k}(t) \leq u|\mathcal{F}_{t-1}]du$$

Under the one-dimensional i.i.d. mobility model, we have,

$$\mathbf{P}[\widetilde{L}_{p,k}(t) \leq u|\mathcal{F}_{t-1}] \leq 1 - (1 - 2u)^{r_p(t)} \leq 2r_p(t)u$$

Therefore,

$$\mathbb{E}\left[\frac{1}{L_{p,k}(t)}|\mathcal{F}_{t-1}\right] \leq 8r_p(t)\log n.$$

when $n \geq 3$. Finally, since $r_p(t)$ is $\mathcal{F}_{t-1}$-measurable, we have

$$\mathbb{E}\left[\frac{1}{L_{p,k}(t)r_p(t)}|\mathcal{F}_{t-1}\right] = \frac{1}{r_p(t)}\mathbb{E}\left[\frac{1}{L_{p,k}(t)}|\mathcal{F}_{t-1}\right] \leq 8\log n.$$

$\square$

Proof of *Lemma 6*: Let

$$V_t = 8\log n[t - t_0(p)] - \sum_{s=t_0(p)+1}^{t} \frac{1}{L_{p,k}(t)r_p(t)}\mathbf{I}_{\{C_{p,k}(t)=1\}},$$

where $C_{p,k}(t) = 1$ denotes that the scheduler decides that a successful capture of packet p to destination k occurs at time-slot t. Then for all $t \geq t_0(p)$, V_t is also $\mathcal{F}_t$-measurable and $V_{t_0(p)} = 0$. By *Lemma 7*, we have

$$\mathbb{E}[V_t - V_{t-1}|\mathcal{F}_{t-1}]$$
$$= 8\log n - \mathbb{E}\left[\frac{1}{L_{p,k}(t,k)r_p(t)}\mathbf{I}_{\{C_{p,k}(t)=1\}}\right]$$
$$\geq 8\log n - \mathbb{E}\left[\frac{1}{L_{p,k}(t)r_{p,k}(t)}|\mathcal{F}_{t-1}\right]$$
$$\geq 0.$$

Hence,

$$\mathbb{E}[V_t|\mathcal{F}_{t-1}] \geq V_{t-1}$$

i.e., V_t is a sub-martingale. Let $s_{p,k} = \min\{t : t \geq t_0(p) \text{ and } C_{p,k}(t) = 1\}$. Since $s_{p,k}$ is a stopping time, by appropriately invoking the Optional Stopping Theorem (Theorem 4.1 in [23]), we have,

$$\mathbb{E}[V_{s_{p,k}}] \geq 0$$

Hence,

$$8\log n\mathbb{E}[D_{p,k}] \geq \mathbb{E}\left[\frac{1}{L_{p,k}(s_{p,k})R_{p,k}}\right].$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON MOBILE COMPUTING

10

Using Hölder's Inequality,

$$\mathbb{E}\left[\frac{1}{L_{p,k}(s_{p,k})}\right] \leq \mathbb{E}[R_{p,k}]\mathbb{E}\left[\frac{1}{L_{p,k}(s_{p,k})R_{p,k}}\right].$$

Then we have

$$8\log n\mathbb{E}[D_{p,k}] \geq \mathbb{E}\left[\frac{1}{L_{p,k}(s_{p,k})}\right]\frac{1}{\mathbb{E}[R_{p,k}]} \geq \frac{1}{\mathbb{E}[L_{p,k}(s_{p,k})]\mathbb{E}[R_{p,k}]}.$$

Finally, by definition,

$$l_{p,k} = \widetilde{l}_{p,k}(s_{p,k}) \geq L_{p,k}(s_{p,k}) - \frac{1}{n}$$

Therefore,

$$8\log n\mathbb{E}[D_{p,k}] \geq \frac{1}{(\mathbb{E}[l_{p,k}] + \frac{1}{n})\mathbb{E}[R_{p,k}]}$$

Thus Lemma 6 holds,

$$c_6\log n\mathbb{E}[D_p] \geq \frac{1}{\left(\mathbb{E}[L_p] + \frac{1}{n}\right)\mathbb{E}[R_p]}. \tag{21}$$

$\square$

By *Lemma 2* and *Lemma 6*, we have the following theorem.

Theorem 5: Under *one-dimensional i.i.d. fast mobility model*, let D be the mean delay averaged over all packets, and let λ be the capacity per multicast session. When $D = o\left(\frac{\sqrt{n}}{n_d}\right)$, the following upper bound holds for any causal scheduling policy,

$$\lambda = O\left(\frac{n}{n_s n_d}\sqrt[3]{\frac{n_d^2 D^2}{n}\log^3 n}\right). \tag{22}$$

6.2 Achievable Lower Bound

We first present the optimal values of key parameters in one-dimensional i.i.d. fast mobility model in Table 3.

TABLE 3
The order of the optimal values of the parameters in one-dimensional fast i.i.d. mobility model.

L: Capture Range	$\Theta(n^{-\frac{1+\alpha+d}{3}}/\log^{\frac{1}{3}} n)$
R: # of Duplicates	$\Theta(n^{\frac{1+\alpha-2d}{3}}/\log^{\frac{2}{3}} n)$

Capacity Achieving Scheme III:

We propose a flexible *rectangle-partition* scheme, similar to [5], to achieve a capacity-delay tradeoff that is close to the upper bound. Rectangle-partition model divides the unit square into multiple horizontal rectangles, named as H-rectangles; and multiple vertical rectangles, named as V-rectangles as in Figure 2. A packet is said to be destined to a rectangle if the orbit of one of its destinations is contained in the rectangle. Each H-rectangle and V-rectangle cross to form a cell, and transmissions only happen in the same *crossing cell*. The transmission of a packet in the H(V) multicast session will go
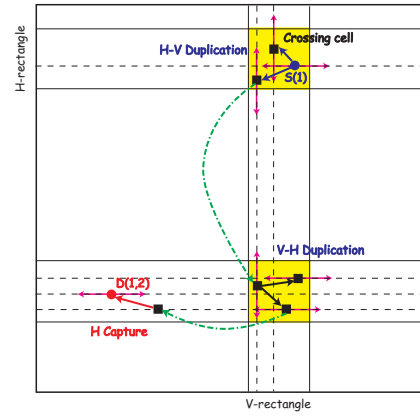

Fig. 2. One-dimensional transmissions in Scheme III.

through H(V)-V(H) duplication, V(H)-H(V) duplication and H(V)-H(V) capture, three procedures, sequentially (see Figure 2).

We group every D time-slots into a super-slot, and let z denote any non-negative integer.

1) At each $3z + 1$ super-slot, we schedule transmissions from the H(V)-sources to the V(H)-relays in every time-slot. We divide the unit square into $\mathcal{R}_d$ H-rectangles and $\mathcal{R}_d$ V-rectangles, i.e., $\mathcal{R}_d^2 = \Theta\left(\frac{n^{(2-\alpha+2d)/3}}{\log n}\right)$ crossing cells. Each cell is a square of area $1/\mathcal{R}_d^2$. We refer to each cell in the $3z + 1$ super-slot as a *duplication cell*. By *Lemma 6* in [4], each cell can be active for $1/c_4$ amount of time, where c_4 is some constant. When a cell is scheduled to be active, each H(V)-source node in the cell broadcasts a new packet to all other V(H)-nodes in the same cell for $\Theta\left(\frac{n^{-(3s+\alpha-2-2d)/3}}{\log^2 n}\right)$ amount of time. These other V(H)-nodes then serve as mobile V(H)-relays for the packet to complete the *V(H)-H(V) duplications* in the next super-slot. The source nodes within the same *duplication cell* coordinate themselves to broadcast sequentially.

2) At each $3z + 2$ super-slot, we schedule transmissions from the V(H)-relay nodes to the H(V)-relay nodes in every time-slot. We use the same partition method as the one used in $3z + 1$ super-slot. When a cell is scheduled to be active, search for V(H)-relay nodes holding the packet, which is destined to the H(V)-rectangle containing this crossing cell and has not been *V(H)-H(V) duplicated* yet. If there are multiple satisfied V(H)-nodes for one packet, randomly choose one and broadcast the packet to all other H(V)-nodes in the same cell. We can easily prove that with R V(H)-relay nodes for each packet p, which are generated in *H(V)-V(H) duplication* of former $3z + 1$ super-slot, w.h.p., there must be a time-slot within this $3z + 2$ super-slot that at least one of them reaches the destined H(V)-rectangle of packet p. And under proper scheduling, the throughput in this period cannot be smaller than that in $3z + 1$ super-slot.

3) At each $3z + 3$ super-slot, we schedule transmissions from the mobile H(V)-relays to the H(V)-destination nodes in every time-slot. We divide the unit square into $\mathcal{R}_c = \Theta\left(n^{(1+\alpha+d)/3}\right)$ H-rectangles and $\mathcal{R}_c$ V-rectangles, i.e., $\mathcal{R}_c^2$ crossing cells. Each cell is a square of area $1/\mathcal{R}_c^2$. We refer to each cell in the $3z+3$ super-slot as the *capture cell*. In each time-slot, for each H(V)-destination node $\mathcal{D}$ and each of its destined packet p, search for H(V)-relay nodes in the same *capture cell* holding packet p. If there are multiple ones, randomly pick one, which we call a *representative H(V)-relay*, and transmit the destined packet p to $\mathcal{D}$. In the end of each $3z + 3$ super-slot, clear all the buffers of mobile nodes, and prepare for a new turn of duplications and capture.

Following the proof of Proposition 2, we have

Proposition 3: With probability approaching one, as $n \to \infty$, the above scheme allows each source to send D packets of size $\lambda = \Theta\left(\frac{n^{-(3s+\alpha-2-2d)/3}}{\log^2 n}\right)$ to their respective destinations within $3D$ time-slots.

7 ONE DIMENSIONAL I.I.D. SLOW MOBILITY MODEL

In this section, we study the one-dimensional i.i.d. slow mobility model.

7.1 Upper Bound

By *Lemma 3* and *Lemma 6*, we have the following theorem.

Theorem 6: Under *one-dimensional i.i.d. slow mobility model*, let D be the mean delay averaged over all packets, and let λ be the capacity per multicast session. When $D = o\left(\frac{\sqrt{n}}{n_d}\right)$, the following upper bound holds for any causal scheduling policy,

$$\lambda = O\left(\frac{n}{n_s n_d} \sqrt[4]{\frac{n_d^2 D^2}{n} \log^3 n}\right). \tag{23}$$

7.2 Achievable Lower Bound

We first present the optimal values of key parameters in one-dimensional i.i.d. slow mobility model in Table 4.

TABLE 4
The order of the optimal values of the parameters in one-dimensional slow i.i.d. mobility model.

L: Capture Range	$\Theta(n^{-\frac{1+2\alpha+2d}{4}}/\log^{\frac{3}{4}} n)$
R: # of Duplicates	$\Theta(n^{\frac{1+2\alpha-2d}{4}}/\log^{\frac{1}{4}} n)$
H: # of Hops	$\Theta(n^{\frac{1-2\alpha-2d}{4}}/\log^{\frac{5}{4}} n)$
S: Hop Length	$\Theta(\sqrt{\log n/n})$

Capacity Achieving Scheme IV: We group every D time-slots into a super-slot, and let z denote any non-negative integer. Scheme IV is similar to Capacity

Achieving Scheme III, presented in Section 6.2, and we only introduce the differences here.

1) At each $3z + 1$ super-slot, we schedule transmissions from the H(V)-sources to the V(H)-relays in every time-slot. We divide the unit square into $\mathcal{R}_d$ H-rectangles and $\mathcal{R}_d$ V-rectangles, i.e., $\mathcal{R}_d^2 = \Theta\left(\frac{n^{(3-2\alpha+2d)/4}}{\log n}\right)$ crossing cells. When a cell is scheduled to be active, each H(V)-source node in the cell broadcasts a new packet to all other V(H)-nodes in the same cell for $\Theta\left(\frac{n^{-(4s+2\alpha-3-2d)/4}}{\log^2 n}\right)$ amount of time.

2) The same as Capacity Achieving Scheme III (2).

3) At each $3z + 3$ super-slot, we schedule transmissions from the mobile H(V)-relays to the H(V)-destination nodes in every time-slot. We divide the unit square into $\mathcal{R}_c = \Theta\left(n^{(1+2\alpha+2d)/4}\right)$ H-rectangles and $\mathcal{R}_c$ V-rectangles, i.e., $\mathcal{R}_c^2$ crossing cells. After picking out a representative H(V)-relay, we then schedule multi-hop transmissions in the following fashion to forward this destined packet p from the *representative H(V)-relay* to $\mathcal{D}$. We further divide each *capture cell* into $\mathcal{R}_h = \Theta\left(\frac{n^{(1-2\alpha-2d)/4}}{\sqrt{\log n}}\right)$ H-rectangles and $\mathcal{R}_h$ V-rectangles, i.e., $\mathcal{R}_h^2$ crossing *hop-cells*. Each *hop-cell* is a square of side length $1/(\mathcal{R}_c \mathcal{R}_h)$. By *Lemma 6* in [4], there exists a scheduling scheme where each hop-cell can be active for $1/c_4$ amount of time. When each hop-cell is active, it forwards a packet to another H(V)-node in the neighboring *hop-cell*. If the H(V)-destination node of the packet is in the neighboring cell, the packet is forwarded directly to the H(V)-destination node. The packets from each *representative H(V)-relay* are first forwarded towards neighboring cells along the X-axis, then to their destination nodes along the Y-axis. At the end of each $3z + 3$ super-slot, clear all the buffers of mobile nodes, and prepare for a new turn of duplications and capture.

Proposition 4: With probability approaching one, as $n \to \infty$, the above scheme allows each source to send D packets of size $\lambda = \Theta\left(\frac{n^{-(4s+2\alpha-3-2d)/4}}{\log^2 n}\right)$ to their respective destinations within $3D$ time-slots.

8 ONE DIMENSIONAL HYBRID R.W. MOBILITY MODEL

In this section, we present the optimal multicast capacity-delay tradeoffs of the one-dimensional hybrid random walk mobility model with both fast and slow mobiles. The results can be established by following similar analysis as one-dimensional i.i.d. mobility models. The details are omitted here for brevity.

Theorem 7: Under *one-dimensional hybrid random walk fast mobility model*, let D be the mean delay averaged over all packets, and let λ be the capacity per multicast session. When $B = o(1)$ and $D = \omega(1/B^2)$, the following

upper bound holds for any causal scheduling policy,

$$\lambda = O\left(\frac{n}{n_s n_d} \sqrt[3]{\frac{n_d^2 D^2}{n} \log^3 n}\right). \tag{24}$$

Theorem 8: Under *one-dimensional hybrid random walk slow mobility model*, let D be the mean delay averaged over all packets, and let λ be the capacity per multicast session. When $B = o(1)$ and $D = \omega(1/B^2)$, the following upper bound holds for any causal scheduling policy,

$$\lambda = O\left(\frac{n}{n_s n_d} \sqrt[4]{\frac{n_d^2 D^2}{n} \log^3 n}\right). \tag{25}$$

9 RESULTS DISCUSSIONS

Our results of optimal multicast capacity-delay tradeoffs in mobile ad hoc networks give a more general result than previous works:

- It generalizes the optimal delay-throughput trade-offs in unicast traffic pattern in [5], when we set $n_s = n$ and $n_d = 1$.
- It generalizes the multicast capacity result $O(\sqrt{D/n_s})$ under delay constraint in [15], which considers the two-dimensional i.i.d. fast mobility model and provides better results than [14], when we set $n_s n_d = n$.

We summarize our results in Table 5 where we omit the logarithmic factors. Setting $n_s = n$ and $n_d = 1$, our results are shown in the second column. Setting $n_s = n$ and $n_d = k$, our results are shown in the third column.

TABLE 5
Optimal multicast capacity and delay tradeoffs in MANETs: a global perspective

λ (i.i.d./hybrid r.w.)	unicast	multicast
2D fast mobility	$O\left(\sqrt{\frac{D}{n}}\right)$	$O\left(\frac{1}{k}\sqrt{\frac{D}{n}k}\right)$
2D slow mobility	$O\left(\sqrt[3]{\frac{D}{n}}\right)$	$O\left(\frac{1}{k}\sqrt[3]{\frac{D}{n}k}\right)$
1D fast mobility	$O\left(\sqrt[3]{\frac{D^2}{n}}\right)$	$O\left(\frac{1}{k}\sqrt[3]{\frac{D^2}{n}k^2}\right)$
1D slow mobility	$O\left(\sqrt[4]{\frac{D^2}{n}}\right)$	$O\left(\frac{1}{k}\sqrt[4]{\frac{D^2}{n}k^2}\right)$

We would like to mention that, similar to the unicast case, [5], our one-dimensional mobility models achieve a larger capacity than two-dimensional models under the multicast traffic pattern. The advantage of lower dimensional mobility lies in the fact that it is simple and easily predictable, thus increasing the inter contact rate. Though nodes are limited to only moving horizontally or vertically, the mobility range on their orbit lines is not restricted. Moreover, for H(V) multicast sessions, the V(H)-relay nodes are used to compensate for the lack of vertical(horizontal) mobility. Given the above analysis, the one-dimensional mobility model in our paper is actually a *hybrid dimensional model*, where

one-dimensional mobile nodes transmit packets in two-dimensional space. We plan to study the capacity improvement brought about by this hybrid dimensional model in the future.

10 HETEROGENEOUS NETWORKS WITH IN-FRASTRUCTURE SUPPORT

In heterogeneous networks, transmissions can be carried out either in infrastructure mode or in ad hoc mode (see Figure 3). Let T_i and T_a denote the maximum *aggregate input throughput* of the whole network when all the transmissions are carried out in infrastructure mode and in ad hoc mode, respectively. Then, the aggregate input capacity of heterogeneous wireless networks, denoted by T, can be calculated as follows:

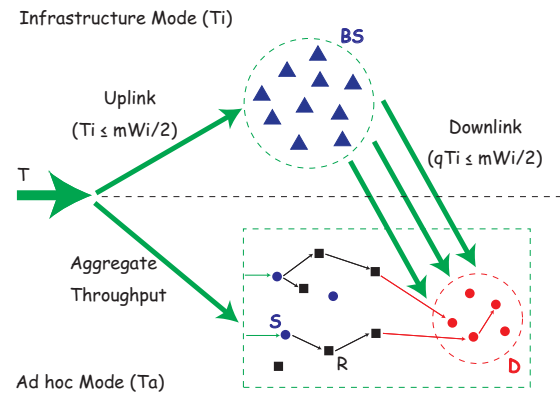$$T = \max\{T_i, T_a\}. \tag{26}$$

Fig. 3. Two modes in heterogeneous networks.

T_a has been studied in previous sections under different mobility models. We only discuss T_i here. Since all base stations share information simultaneously, they can be regarded as an integrated relay. We define this specific huge relay as *BS relay*, with the same maximum input and output throughput $mW_i/2$. Please recall that $m = n^\beta$ is the number of base stations in the network.

Under infrastructure mode, packets should be transmitted by three steps. Packets is sent from source to the nearest base station first and then shared by all the base stations. At the third step, packets are sent to the destination by the nearest base station. Since we assume the base stations are wired connected, the bandwidth can be regarded as large enough and the delay at the second step is ignored.

To conduct further analysis, we present a lemma first to bound the number of nodes in a given region.

Lemma 8: For each packet and in each subregion, w.h.p., there are at most $\Theta(\frac{n_d}{m})$ destination nodes when $m = o(n_d)$, and at most $\Theta(1)$ destination nodes when $m = \omega(n_d)$.

Proof: Consider subregion i. Let X_i be a random variable that denotes the number of destination nodes

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON MOBILE COMPUTING

13

in subregion i. Then we have $\mathbb{E}[X_i] = \frac{n_d}{m}$. Recall the Chernoff Bound in [24], for any $\delta > 0$,

$$Pr\big(X_i > (1+\delta)\mathbb{E}[X_i]\big) < e^{-\mathbb{E}[X_i]f(\delta)} \qquad (27)$$

where $f(\delta) = (1+\delta)\log(1+\delta) - \delta$.

1) $m = o(n_d)$, i.e., $0 \le \beta < \alpha \le 1$.

According to the Chernoff bound (27), we have

$$Pr(X_i > 2\frac{n_d}{m}) < e^{-\frac{n_d}{m}f(1)}$$

where $f(1) = 2\log 2 - 1 > 0$. Since $0 \le \beta < \alpha$, when n is large enough, we have

$$
\begin{aligned}
Pr(X_i \le 2\frac{n_d}{m} \text{ for any i }) &\ge 1 - mPr(X_i > 2\frac{n_d}{m}) \\
&> 1 - n^\beta e^{-n^{\alpha-\beta}f(1)} \\
&\to 1
\end{aligned}
$$

2) $m = \omega(n_d)$, i.e., $0 \le \alpha < \beta \le 1$.

According to the Chernoff bound (27), we have

$$Pr(X_i > (1+\delta)\frac{n_d}{m}) < \frac{e^{\delta\frac{n_d}{m}}}{(1+\delta)^{(1+\delta)\frac{n_d}{m}}}.$$

Let $(1+\delta)\frac{n_d}{m} = c_7$, where c_7 is a constant that will be determined later. Then we have

$$Pr(X_i > c_7) < \frac{e^{2-n^{\alpha-\beta}}}{c_7 n^{c_7(\beta-\alpha)}}.$$

Hence,

$$Pr(X_i \le c_7 \text{ for any i }) > 1 - \frac{e^{2-n^{\alpha-\beta}}}{c_7 n^{(c_7-1)\beta-c_7\alpha}}.$$

We can choose $c_7 > \frac{\beta}{\beta-\alpha}$, so that $(c_7-1)\beta - c_7\alpha > 0$. This probability goes to 1 as n goes to infinity.
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

For the first step, when $m = \Theta(n_s)$, there will be at most $\log n$ source nodes inside every subregion. Therefore, the aggregate input throughput for the first step is $T_{i1} = \Theta(\frac{m}{\log n})$. When $m = \omega(n_s)$, there will be at most $\Theta(1)$ source nodes inside every subregion and $T_{i1} = \Theta(n_s)$. When $m = o(n_s)$, there will be at least $\Theta(\frac{n_s}{m})$ source nodes inside every subregion and $T_{i1} = \Theta(m)$.

Under multicast traffic pattern, multiple deliveries are needed to reach all the n_d destinations. Hence, for each input flow of the *BS relay*, there are at least q output flows. Note that $q \ne n_d$, for there may be multiple destination nodes within the same subregion and only one broadcast is enough for all of them.

For the third step, if $m = \omega(n_d)$, there will be at most constant destinations in every subregion for each multicast session. Since $n_s n_d \ge n \ge m$, there will be at least $\Theta(\frac{n_s n_d}{m})$ output flows for each base station. Hence, the aggregate input throughput for the third step is $T_{i2} = \Theta(\frac{m^2}{n_s n_d})$. When $m = o(n_d)$, there are at least $\Theta(\frac{n_d}{m})$ destinations for every multicast session in every subregion, which means that the base station should broadcast the information for every multicast session and therefore $T_{i2} = \Theta(m)$.

Combined the above results together, we can obtain the aggregate input throughput for the network. It can also be inferred that base stations may not help unless there are a large number of them.

11 CONCLUSION

In this paper, we have studied the multicast capacity-delay tradeoffs in both homogeneous and heterogeneous mobile networks. Specifically, in homogeneous networks, we established the upper bound on the optimal multicast capacity-delay tradeoffs under two-dimensional/one-dimensional i.i.d./hybrid random walk fast/slow mobility models and proposed capacity achieving schemes to achieve capacity close to the upper bound. In addition, we find that though the one-dimensional mobility model constrains the direction of nodes' mobility, it achieves larger capacity than the two-dimensional model since it is more predictable. Also, slow mobility brings better performance than fast mobility because there are more possible routing schemes.

12 ACKNOWLEDGMENT

REFERENCES

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

[2] X. Lin, G. Sharma, R. R. Mazumdar and N. B. Shroff, "Degenerate delay-capacity tradeoffs in ad-hoc networks with Brownian mobility," in *IEEE Trans. on Inform. Theory*, vol.52, no. 6, pp. 277-2784, June 2006.

[3] M. Neely and E. Modiano, "Capacity and Delay Tradeoffs for Ad-hoc Mobile Networks," in *IEEE Trans. on Inform. Theory*, vol. 51, no. 6, pp. 1917-1937, 2005.

[4] X. Lin and N. B. Shroff, "The Fundamental Capacity-Delay Tradeoff in Large Mobile Ad Hoc Networks" in *Proc. Third Annu. Mediterranean Ad Hoc Netw. Workshop*, 2004.

[5] L. Ying, S. Yang and R. Srikant, "Optimal delay-throughput tradeoffs in mobile ad-hoc networks," in *IEEE Trans. on Inform. Theory*, vol. 9, no. 54, pp. 4119-4143, September 2008.

[6] J. Mammen and D. Shah, "Throughput and delay in random wireless networks with restricted mobility," in *IEEE Trans. on Inform. Theory*, vol. 53, no. 3, pp. 1108-1116, March 2007.

[7] P. Li, Y. Fang and J. Li, "Throughput, Delay, and Mobility in Wireless Ad Hoc Networks," in *Proc. IEEE INFOCOM*, San Diego, CA, March 15-19, 2010.

[8] U. Kozat and L. Tassiulas, "Throughput Capacity of Random Ad Hoc Networks with Infrastructure Support." in *Proc. ACM MobiCom*, San Diego, CA, USA, June 2003.

[9] P. Li, C. Zhang and Y. Fang, "Capacity and Delay of Hybrid Wireless Broadband Access Networks." in *IEEE Journal on Selected Areas in Communications (JSAC)* - Special Issue on Broadband Access Networks, 27(2):117-125, February 2009.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

IEEE TRANSACTIONS ON MOBILE COMPUTING

14

[10] B. Liu, Z. Liu and D. Towsley, "On the Capacity of Hybrid Wireless Networks." in *Proc. IEEE INFOCOM*, San Francisco, California, USA, March 2003.

[11] X. Li, S. Tang and O. Frieder, "Multicast capacity for large scale wireless ad hoc networks," in *Proc. ACM MobiCom*, Sept. 2007.

[12] S. Shakkottai, X. Liu and R. Srikant, "The multicast capacity of large multihop wireless networks," in *Proc. ACM MobiHoc*, Sept. 2007.

[13] U. Niesen, P. Gupta and D. Shah, "The multicast capacity region of large wireless networks," in *Proc. IEEE INFOCOM*, Apr. 2009.

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

[15] S. Zhou and Lei Ying, "On Delay Constrained Multicast Capacity of Large-Scale Mobile Ad-Hoc Networks." in *Proc. INFOCOM 2010 mini-conference*, San Diego, CA, 2010.

[16] P. Li and Y. Fang, "Impacts of Topology and Traffic Pattern on Capacity of Hybrid Wireless Networks." in *IEEE Trans. on Mobile Computing*, vol. 8, no. 12, pp: 1585 - 1595, Dec. 2009.

[17] X. Mao, X. Li and S. Tang, "Multicast Capacity for Hybrid Wireless Networks." in *ACM MobiHoc*, 2008.

[18] P. Li and Y. Fang, "The Capacity of Heterogeneous Wireless Networks," in *Proc. IEEE INFOCOM*, San Diego, CA, March 15-19, 2010.

[19] Z. Wang, S. Karande, H. Sadjadpour and J. Garcia-Luna-Aceves, "On the Multicast Capacity of Wireless Ad Hoc Vetworks with Network Coding." in *IEEE Journal of Communications and Networks*, vol. 13, no. 5, pp: 525 - 535, Oct. 2011.

[20] C. Wang, C. Jiang, X-Y. Li, S. Tang, Y. He, X. Mao, and Y. Liu , "Scaling Laws of Multicast Capacity for Power-Constrained Wireless Networks under Gaussian Channel Model." in *IEEE Trans. on Computers*, vol. 61, no. 5, pp: 713 - 725, May 2012.

[21] P. Li, Y. Fang, J. Li, and X. Huang, "Smooth Trade-offs Between Throughput and Delay in Mobile Ad Hoc Networks." in *IEEE Trans. on Mobile Computing*, vol. 11, no. 3, pp. 427-438, March 2012.

[22] C. Wang, S. Tang, X-Y. Li, and C. Jiang, "Multicast Capacity Scaling Laws for Multihop Cognitive Networks." in *IEEE Trans. on Mobile Computing*, vol. PP, no. 99, pp: 1 - 1, 2012.

[23] R. Durrett, *Probability: Theory and Examples*, 2nd ed. Belmont, CA: Duxbury Press, 1996.

[24] T. Cormen, C. Leiserson, R. Rivest and C. Stein, *Introduction to Algorithms (2nd ed.)*, MIT Press, 2001.

Xiaohua Tian received his B.E. and M.E. degrees in communication engineering from Northwestern Polytechnical University, Xi'an, China, in 2003 and 2006, respectively. He received the Ph.D. degree in the Department of Electrical and Computer Engineering (ECE), Illinois Institute of Technology (IIT), Chicago, in Dec. 2010. He is currently a research associate in Department of Electronic Engineering in Shanghai Jiaotong University, China. His research interests include application-oriented networking, Internet of Things and wireless networks. He serves as the guest editor of International Journal of Sensor Networks, publicity co-chair of the 7th International Conference on Wireless Algorithms, Systems, and Applications (WASA 2012). He also serves as the Technical Program Committee member for Communications QoS, Reliability, and Modeling Symposium (CQRM) of GLOBECOM 2011, Wireless Networking of GLOBECOM 2013 and WASA 2011.

Y un Wang received the BE degree in electronic engineering in Shanghai Jiao Tong University, China in 2011. His research interests are in the area of asymptotic analysis of multicast capacity in wireless ad hoc network.

X iaoyu Chu received the BE degree in electronic engineering in Shanghai Jiao Tong University, China in 2010. Her research interests are in the area of asymptotic analysis of multicast capacity in wireless ad hoc network.

Jinbei Zhang received his B. E. degree in Electronic Engineering from Xidian University, Xi'an, China, in 2010, and is currently pursuing the Ph.D. degree in electronic engineering at Shanghai Jiao Tong University, Shanghai, China.

His current research interests include network security, capacity scaling law and mobility models in wireless networks.
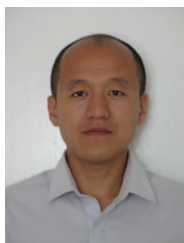
Xinbing Wang received the B.S. degree (with hons.) from the Department of Automation, Shanghai Jiaotong University, Shanghai, China, in 1998, and the M.S. degree from the Department of Computer Science and Technology, Tsinghua University, Beijing, China, in 2001. He received the Ph.D. degree, major in the Department of electrical and Computer Engineering, minor in the Department of Mathematics, North Carolina State University, Raleigh, in 2006. Currently, he is a faculty member in the Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai, China. His research interests include scaling law of wireless networks and cognitive radio. Dr. Wang has been an associate editor for IEEE Transactions on Mobile Computing, and the member of the Technical Program Committees of several conferences including ACM MobiCom 2012, ACM MobiHoc 2012, IEEE INFOCOM 2009-2013.

Yu Cheng received the B.E. and M.E. degrees in Electrical Engineering from Tsinghua University, Beijing, China, in 1995 and 1998, respectively, and the Ph.D. degree in Electrical and Computer Engineering from the University of Waterloo, Waterloo, Ontario, Canada, in 2003. From September 2004 to July 2006, he was a postdoctoral research fellow in the Department of Electrical and Computer Engineering, University of Toronto, Ontario, Canada. Since August 2006, he has been with the Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, Illinois, USA, as an Assistant Professor. His research interests include next-generation Internet architectures and management, wireless network performance analysis, network security, and wireless/wireline interworking. He received a Postdoctoral Fellowship Award from the Natural Sciences and Engineering Research Council of Canada (NSERC) in 2004, and a Best Paper Award from the conferences QShine 2007 and ICC 2011. He received the National Science Foundation (NSF) CAREER award in 2011. He served as a Co-Chair for the Wireless Networking Symposium of IEEE ICC 2009, a Co-Chair for the Communications QoS, Reliability, and Modeling Symposium of IEEE GLOBECOM 2011, and a Technical Program Committee (TPC) Co-Chair for WASA 2011. He is an Associated Editor for IEEE Transactions on Vehicular Technology.