This paper was presented as part of the main technical program at IEEE INFOCOM 2011

Optimal Multicast Capacity and Delay Tradeoffs in MANETs: A Global Perspective

Yun Wang, Xiaoyu Chu, Xinbing Wang
Dept of Electronic Engineering
Shanghai Jiao Tong University, China
Email: {sunshinehk,cxygrace,xwang8}@sjtu.edu.cn

Yu Cheng
Dept. of Elec. and Comp. Engin.
Illinois Institute of Technology, USA
Email: cheng@iit.edu

***Abstract*—In this paper, we give a global perspective of multicast capacity and delay analysis in Mobile Ad-hoc Networks (MANETs). Specifically, we consider two node mobility models: (1) two-dimensional i.i.d. mobility, (2) one-dimensional i.i.d. mobility. Two mobility time-scales are included 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 four 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. Our study can be further extended to two-dimensional/one-dimensional hybrid random walk fast/slow mobility models and heterogeneous networks.**

I. 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 done 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].

All the above works studied the unicast traffic. 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], 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 the number of destinations per source. In their network, $\Theta(n)$ cell-partitioning [3] was used, which fixed the transmission range as $O(\frac{1}{n})$. 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 $O\Big(\min\Big\{1, (\log p)(\log(n_s p))\sqrt{\frac{D}{n_s}}\Big\}\Big)$. Then a *joint coding/scheduling algorithm* was proposed to achieve a throughput of $O\Big(\min\Big\{1, \sqrt{\frac{D}{n_s}}\Big\}\Big)$. 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. [20] 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 [20] all study static networks.

In this paper, we give a general analysis on the optimal multicast capacity-delay tradeoffs in homogeneous MANETs. Our results will be used in our future work to study 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.

We summarize our main results here:

(1) Two-dimensional i.i.d. mobility models:
   - (i) Under the fast mobility assumption, it is shown that the maximum throughput per multicast session is $O\Big(\frac{n}{n_s n_d}\sqrt{\frac{D}{n}n_d}\Big)$ under a delay constraint D. A cell-partitioned scheme is presented to achieve a close capacity when $D = o(\frac{n}{n_d})$ and $n_s n_d \geq n$.
   - (ii) Under the slow mobility assumption, it is shown that the maximum throughput per multicast session is $O\Big(\frac{n}{n_s n_d}\sqrt[3]{\frac{D}{n}n_d}\Big)$ under a delay constraint D. A cell-partitioned scheme is presented to achieve a close capacity when $D = o(\frac{n}{n_d})$ and $n_s n_d \geq n$.

(2) One-dimensional i.i.d. mobility models:
   - (i) Under the fast mobility assumption, it is shown that the maximum throughput per multicast session is $O\Big(\frac{n}{n_s n_d}\sqrt[3]{\frac{D^2}{n}n_d^2}\Big)$ under a delay constraint D. A cell-partitioned scheme is presented to achieve a close capacity when $D = o(\frac{\sqrt{n}}{n_d})$, $n_d = O(\sqrt{n})$ and

$n_s n_d \geq n$.

(ii) Under the slow mobility assumption, it is shown that the maximum throughput per multicast session is $O\left(\frac{n}{n_s n_d}\sqrt[4]{\frac{D^2}{n}n_d^2}\right)$ under a delay constraint D. A cell-partitioned scheme is presented to achieve a close capacity when $D = o(\frac{\sqrt{n}}{n_d})$, $n_d = O(\sqrt{n})$ and $n_s n_d \geq n$.

The rest of the paper is organized as follows. In section II, we outline the system models. Four mobility models are discussed in section III to section VI respectively. Section VII offers some discussion on our results. Then we conclude.

II. System Models

Multicast Traffic Pattern: 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 has n_d distinct destination nodes. We group each source and its n_d destinations as a multicast session. Note that a particular node may be included by different multicast sessions as either source or destination.

Protocol Model: We assume the following Protocol Model from [1] that governs direct radio transmissions between nodes. Let W be the bandwidth of the system. Let X_i denote the position of node $i, i = 1, \ldots, n$. Let $|X_i - X_j|$ be the Euclidean distance between nodes i and j. Node i can communicate directly with another node j at W bits per second if and only if the following interference constraint is satisfied for every other node $k \neq i, j$ that is simultaneously transmitting, [1]. Here, Δ is some positive number.

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

Definition of Capacity: We assume the same packet arrival rate per time-slot for each source, say λ. The network is said stable if and only if there exists a certain scheduling scheme which can guarantee the finite length of queue in each node as time goes to infinity. Then the *capacity*, which is short for per-session capacity, is defined as the maximum arrival rate λ that the stable network can support.

Definition of Delay: We define the *survival time* for a certain packet as the time interval counting from the moment it enters the network and ending until one of its copies reaches the last destination. Note that we consider the expectation value over all possible network configurations. Then the *delay*, which is short for per-session delay, is defined as the average *survival time* over all packets during an enough long term. Also note that we do not consider the queueing delays in the network.

Notations: Given non-negative functions $f(n)$ and $g(n)$:

(1) $f(n) = O(g(n))$ means there exist positive constants c and m such that $f(n) \leq cg(n)$ for all $n \geq m$.
(2) $f(n) = o(g(n))$ means that $\lim_{n\to\infty} f(n)/g(n) = 0$.
(3) $f(n) = \Omega(g(n))$ means there exist positive constants c and m such that $f(n) \geq cg(n)$ for all $n \geq m$.
(4) $f(n) = \omega(g(n))$ means that $\lim_{n\to\infty} g(n)/f(n) = 0$.
(5) $f(n) = \Theta(g(n))$ means that both $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$ hold.

Mobile ad hoc network model: Consider a pure ad hoc network where n wireless mobile nodes are positioned in a unit square. The unit square is assumed to be a torus, where the left and right edges are connected, and top and bottom edges are also connected. We will study the following mobility models, similar to [5], in this paper.

(1) **Two-dimensional i.i.d. mobility model:** Our two-dimensional i.i.d. mobility model is defined as follows:
   (i) At the beginning of each time slot, the nodes are uniformly, randomly positioned in the unit square.
   (ii) The node positions are independent of each other, and independent from time slot to time slot.

(2) **One-dimensional i.i.d. mobility model:** Our one-dimensional i.i.d. mobility model is defined as follows:
   (i) 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.
   (ii) Let (x_i, y_i) denote the position of node i. If node i is an H-node, y_i is fixed and x_i is a value randomly 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.
   (iii) Assume that source and destinations in the same multicast session are the same type of nodes. Also assume that node i is an H-node if i is odd, and a V-node if i is even.
   (iv) The orbit distance of two H(V)-nodes is defined to be the vertical (horizontal) distance of the two nodes.

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-hop transmission is allowed.
- **Slow mobility**: The mobility of nodes is much slower than the transmission of packets, i.e., multi-hop transmissions may happen within a single 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, [4]. We say a packet p is successfully multicast 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 multicast 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 schedule 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 multicast, 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.

III. 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 logarithmic factors.

A. 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, D_p denote the number of time slots it takes to reach the last destination of packet p. And let R_p denote the number of mobile relays holding packet p when the packet reaches its last destination. 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 is some positive constant).

$$c_1 \log n \mathbb{E}[D_p] \geq \frac{1}{\left(\mathbb{E}[L_p] + \frac{1}{n^2}\right)^2 \mathbb{E}[R_p]} \tag{1}$$

Proof: The detailed proof follows similar procedures to Proposition 1 in [4]. We give some intuitive explanation here. Consider a simpler scenario, in which R_p and L_p are constants in different time slots and for different destinations. Then $1-(1-{L_p}^2)^{R_p}$ is the probability that any one out of the R_p nodes can capture destination k in one time slot. It is easy to show that, the average number of time slots needed to capture destination k, is

$$\mathbb{E}[D_{p,k}] = \frac{1}{1-(1-{L_p}^2)^{R_p}} \geq \frac{1}{{L_p}^2 R_p}$$

■

Consider a large enough 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 resources consumption, the following inequality holds for any causal scheduling policy (c_2 is some positive constant).

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

Proof: Consider the Protocol Model, [1]. By the interference constraint, if nodes i and j directly transmit to nodes k and l respectively, at the same time, we have

$$|X_i - X_j| \geq \frac{\Delta}{2}(|X_i - X_k| + |X_j - X_l|)$$

That is, disks of radius $\frac{\Delta}{2}$ times the transmission range centered at the transmitter are disjoint from each other. This property motivates us to measure the radio resources each transmission consumes by the areas of these disjoint disks, [1]. Next we will calculate the radio resources consumption during *Duplication* and *Capture*, respectively.

- *Capture*: For each packet p and each of its destination k, the *one-hop capture*[1] consumes area of $\frac{\pi\Delta^2}{4}(L_{p,k})^2$. Hence, the lower 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 a lower 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,

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

Proof: Since we assume that no node can transmit and receive over the same frequency at the same time, the following property can be shown as in [1]

$$\frac{WT}{2}n \geq \sum_{p=1}^{\lambda n_s T}\sum_{k=1}^{n_d} 1 = \lambda n_s T n_d$$

Hence, $\lambda \leq \frac{Wn}{2n_s n_d} = \Theta(\frac{n}{n_s n_d})$. In addition, each source can send out at most W size of packet per time-slot, i.e., $\lambda \leq$

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

$W = \Theta(1)$. These first two factors are obvious, and we will neglect them in later results.

By *Lemma 1* and *Lemma 2*, using Jensen's Inequality and Cauchy-Schwartz Inequality, we have

$$\frac{4c_2 WT\log n}{\Delta^2} \geq \frac{1}{c_1 n\log n}\frac{(\lambda n_s T)^3}{D\Big(\sum_{p=1}^{\lambda n_s T}\big(\mathbb{E}[L_p]+\frac{1}{n^2}\big)\Big)^2} + \frac{\pi n_d}{\lambda n_s T}\Bigg(\sum_{p=1}^{\lambda n_s T}\mathbb{E}[L_p]\Bigg)^2 - \lambda T\frac{n_s n_d}{n}$$

There are two cases we need to consider.

Case 1: If $\sum_{p=1}^{\lambda n_s T}\mathbb{E}[L_p] \leq \frac{\lambda n_s T}{n^2}$ then,

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

When $D = o(\frac{n}{n_d})$, the first term dominates when n is large. Hence, for n large enough,

$$\lambda \leq \frac{32c_1c_2W}{\Delta^2}\frac{D\log^2 n}{n_s n^3} \tag{4}$$

Case 2: If $\sum_{p=1}^{\lambda n_s T}\mathbb{E}[L_p] \geq \frac{\lambda n_s T}{n^2}$ then,

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

Since $D = o\big(\frac{n}{n_d\log n}\big)$, the first term dominates when n is large, i.e.,

$$\lambda \leq \sqrt{\frac{64c_1c_2^2W^2}{\pi\Delta^4}}\frac{n}{n_s n_d}\sqrt{\frac{D\log^3 n}{n}n_d} \tag{5}$$

Finally, we compare the two inequalities we have obtained, i.e., (4) and (5). Since $D = o\big(\frac{n}{n_d}\big)$ inequality (5) will eventually be the loosest for large n. Hence, the optimal capacity-delay tradeoff is upper bounded by

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

■

B. 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

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

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 the optimal choices of various key parameters of the scheduling policy. We can infer that the parameters (such as $\mathbb{E}[D_{p,k}]$, $\mathbb{E}[L_{p,k}]$) 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 limit the mobile nodes $n \leq n_s n_d$. The results are summarized in Table I.

TABLE I

THE ORDER OF THE OPTIMAL VALUES OF THE PARAMETERS IN TWO-DIMENSIONAL FAST I.I.D. MOBILITY MODEL. ($n_s = n^s$ MULTICAST SESSIONS ARE INCLUDED, EACH OF WHICH HAS $n_d = n^\alpha$ DESTINATIONS. DELAY IS BOUNDED BY $D = n^d$.)

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

Capacity Achieving Scheme I:

We propose a more flexible cell-partitioning scheme, [4], 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 [20] shows, each cell in the network can transmit at a rate of c_3W, 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 $\mathcal{C}_d = \Theta\big(\frac{n^{(1-\alpha+d)/2}}{\log n}\big)$ 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 $\Theta\big(\frac{n^{-(2s+\alpha-1-d)/2}}{\log^2 n}\big)$ 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 $\mathcal{C}_c = \Theta\big(n^{(1+\alpha+d)/2}\big)$ 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 $Y_{\mathcal{SD}}$ 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 $\mathcal{S}$ generated in that very time-slot. 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.

We can show that as $n \to \infty$, with high probability (w.h.p.), all packets generated in odd duplication super-slot will complete n_d transmissions 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 $\lambda = \Theta\big(\frac{n^{-(2s+\alpha-1-d)/2}}{\log^2 n}\big)$ to their respective destinations within $2D$ time-slots.

Proof: Similar but simpler proof of *Proposition 2*. ■

Remarks: Our scheme uses different cell-partitioning in the odd super-slot that that in the even super-slot. The size of the duplication cell is chosen such that the average number of nodes in each cell, $n/\mathcal{C}_d$, is close to the optimal value of R. The size of the capture cell is chosen such that its area, $1/\mathcal{C}_c$ is close to the optimal value of L^2.

IV. 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.

A. 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 takes from the last mobile relay to destination k. And let $S^h_{p,k}$, $h = 1, 2, \dots, h_{p,k}$ 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).

$$\sum_{p=1}^{\lambda n_s T} \frac{\Delta^2}{4} \frac{\mathbb{E}[R_p] - n_d}{n} + \sum_{p=1}^{\lambda n_s T} \sum_{k=1}^{n_d} \sum_{h=1}^{h_{p,k}} \frac{\pi \Delta^2}{4} \mathbb{E}\big[(S^h_{p,k})^2\big] \leq c_5 W T \log n, \quad (6)$$

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:

$$\sum_{h=1}^{h_{p,k}} S^h_{p,k} \geq L_{p,k} \quad (7)$$

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,

$$\lambda = O\left(\frac{n}{n_s n_d} \sqrt[3]{\frac{n_d D}{n}}\right) \quad (8)$$

Proof: By *Lemma 3*, using Jensen's Inequality and Cauchy-Schwartz Inequality, we have

$$\begin{aligned} &\frac{4c_5 W T \log n}{\Delta^2} \\ \geq\ & \frac{1}{c_1 n \log n} \frac{(\lambda n_s T)^3}{D\Big(\sum_{p=1}^{\lambda n_s T}\big(\mathbb{E}[L_p] + \frac{1}{n^2}\big)\Big)^2} \\ &+ \frac{2\pi n_d^2}{W T n}\Big(\sum_{p=1}^{\lambda n_s T} \mathbb{E}[L_p]\Big)^2 - \lambda T \frac{n_s n_d}{n} \end{aligned}$$

Case 1: If $\sum_{p=1}^{\lambda n_s T} \mathbb{E}[L_p] \leq \frac{\lambda n_s T}{n^2}$ then,

$$\begin{aligned} &\frac{4c_5 W T \log n}{\Delta^2} \\ \geq\ & \frac{1}{4c_1 \log n} \frac{\lambda T n_s n^3}{D} - \lambda T \frac{n_s n_d}{n} \end{aligned}$$

When $D = o(\frac{n}{n_d})$, the first term dominates when n is large. Hence, for n large enough,

$$\lambda \leq \frac{32 c_1 c_5 W}{\Delta^2} \frac{D \log^2 n}{n_s n^3} \quad (9)$$

Case 2: If $\sum_{p=1}^{\lambda n_s T} \mathbb{E}[L_p] \geq \frac{\lambda n_s T}{n^2}$ then,

$$\begin{aligned} &\frac{4c_5 W T \log n}{\Delta^2} \\ \geq\ & \lambda T \frac{n_s n_d}{n} \sqrt{\frac{2\pi \lambda n_s}{c_1 W D \log n}} - \lambda T \frac{n_s n_d}{n} \end{aligned}$$

Therefore, since $D = o(\frac{n}{n_d})$, either

$$\lambda \leq \frac{c_1 W}{2\pi} \frac{D \log n}{n_s} \quad (10)$$

Or if $\lambda = \omega(\frac{D \log n}{n_s})$,

$$\lambda \leq \sqrt[3]{\frac{32 c_1 c_5^2 W^3}{\Delta^4}} \frac{n}{n_s n_d} \sqrt[3]{\frac{D \log^3 n}{n} n_d} \quad (11)$$

Finally, we compare the three inequalities we have obtained, i.e., (9), (10) and (11). Since $D = o\big(\frac{n}{n_d}\big)$ inequality (11) will eventually be the loosest for large n. Hence, the optimal capacity-delay tradeoff is upper bounded by

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

■

TABLE II

THE ORDER OF THE OPTIMAL VALUES OF THE PARAMETERS IN TWO-DIMENSIONAL SLOW I.I.D. MOBILITY MODEL. ($n_s = n^s$ MULTICAST SESSIONS ARE INCLUDED, EACH OF WHICH HAS $n_d = n^\alpha$ DESTINATIONS. DELAY IS BOUNDED BY $D = n^d$.)

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	$\Theta(\sqrt{\log n/n})$

B. Achievable Lower Bound

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

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

The idea is similar, as is presented in Section III-B. We summarize the *optimal values* in Table II.

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 III-B, 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 $\mathcal{C}_d = \Theta\big(\frac{n^{(2-2\alpha+d)/3}}{\log n}\big)$ cells. When a cell is scheduled to be active, each node in the cell broadcasts a new packet to all other nodes in the same cell for $\Theta\big(\frac{n^{-(3s+2\alpha-2-d)/3}}{\log^2 n}\big)$ 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 $\mathcal{C}_c = \Theta\big(n^{(1+2\alpha+2d)/3}\big)$ 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 $\mathcal{C}_h = \Theta\big(\frac{n^{(2-2\alpha-2d)/3}}{\log n}\big)$ *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 $\lambda = \Theta\big(\frac{n^{-(3s+2\alpha-2-d)/3}}{\log^2 n}\big)$ to their respective destinations within $2D$ time-slots.

Proof: See Appendix A. ■

Remarks: In our scheme, the size of the hop-cell is chosen such that each hop to the neighboring hop-cell is of length $1/\sqrt{\mathcal{C}_c\mathcal{C}_h}$, which is close to the optimal value of S.

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

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

A. Upper Bound

Lemma 4: 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).

$$c_6 \log n \mathbb{E}[D_p] \geq \frac{1}{(\mathbb{E}[L_p] + \frac{1}{n})\mathbb{E}[R_p]} \tag{12}$$

Proof: Let ρ_p denote the distance from any one mobile node of packet p to one of its destinations in a particular time-slot. Under one-dimensional i.i.d. mobility model, when the orbits of two nodes are vertical to each other, $\rho_p \leq L$ holds only if they are in a square with side length $2L$ as in Figure 1.

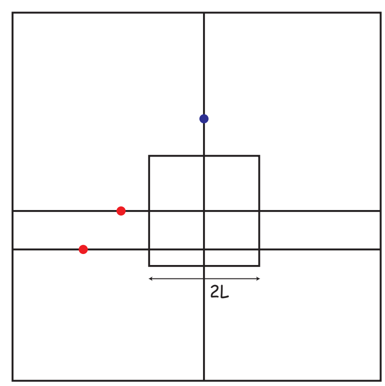

Fig. 1. Vertical and parallel orbits.

In this scenario, we have

$$Pr(\rho_p \leq L) \leq 4L^2$$

When the orbits of these two nodes are parallel to each other, then

$$Pr(\rho_p \leq L) \leq 2L$$

■

By *Lemma 2* and *Lemma 4*, we have the following theorem.

Theorem 3: 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 $D = o\big(\frac{\sqrt{n}}{n_d}\big)$, the following upper bound holds for any causal scheduling policy,

$$\lambda = O\Big(\frac{n}{n_s n_d}\sqrt[3]{\frac{n_d^2 D^2}{n}}\Big) \tag{13}$$

B. Achievable Lower Bound

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

TABLE III
THE ORDER OF THE OPTIMAL VALUES OF THE PARAMETERS IN ONE-DIMENSIONAL FAST I.I.D. MOBILITY MODEL. ($n_s = n^s$ MULTICAST SESSIONS ARE INCLUDED, EACH OF WHICH HAS $n_d = n^\alpha$ DESTINATIONS. DELAY IS BOUNDED BY $D = n^d$.)

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

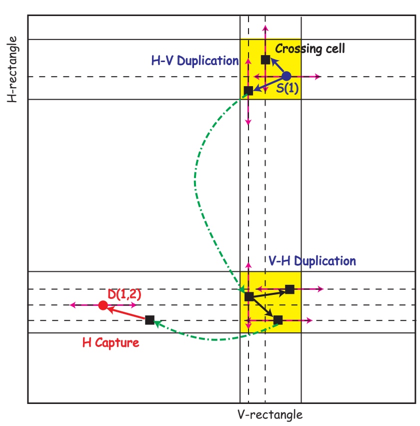

Fig. 2. One-dimensional transmissions in scheme III.

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., $\mathcal{R}_d^2 = \Theta\big(\frac{n^{(2-\alpha+2d)/3}}{\log n}\big)$ 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 $\Theta\big(\frac{n^{-(3s+\alpha-2-2d)/3}}{\log^2 n}\big)$ 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 $\mathcal{R}_c = \Theta\big(n^{(1+\alpha+d)/3}\big)$ 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 $\lambda = \Theta\big(\frac{n^{-(3s+\alpha-2-2d)/3}}{\log^2 n}\big)$ to their respective destinations within $3D$ time-slots.

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

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

A. Upper Bound

By *Lemma 3* and *Lemma 4*, we have the following theorem.

Theorem 4: 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 $D = o\big(\frac{\sqrt{n}}{n_d}\big)$, the following upper bound holds for any causal scheduling policy,

$$\lambda = O\left(\frac{n}{n_s n_d}\sqrt[4]{\frac{n_d^2 D^2}{n}}\right) \tag{14}$$

B. Achievable Lower Bound

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

TABLE IV
THE ORDER OF THE OPTIMAL VALUES OF THE PARAMETERS IN ONE-DIMENSIONAL SLOW I.I.D. MOBILITY MODEL. ($n_s = n^s$ MULTICAST SESSIONS ARE INCLUDED, EACH OF WHICH HAS $n_d = n^\alpha$ DESTINATIONS. DELAY IS BOUNDED BY $D = n^d$.)

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 V-B, 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., $\mathcal{R}_d^2 = \Theta\big(\frac{n^{(3-2\alpha+2d)/4}}{\log n}\big)$ 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 $\Theta\big(\frac{n^{-(4s+2\alpha-3-2d)/4}}{\log^2 n}\big)$ 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 $\mathcal{R}_c = \Theta\big(n^{(1+2\alpha+2d)/4}\big)$ 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 $\mathcal{R}_h = \Theta\big(\frac{n^{(1-2\alpha-2d)/4}}{\sqrt{\log n}}\big)$ 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 $\lambda = \Theta\big(\frac{n^{-(4s+2\alpha-3-2d)/4}}{\log^2 n}\big)$ to their respective destinations within $3D$ time-slots.

VII. RESULTS DISCUSSIONS

Our results of optimal multicast capacity-delay tradeoffs in mobile ad-hoc networks give a global perspective for the following reasons:

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

We summarize our results in Table V. 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 V
OPTIMAL MULTICAST CAPACITY AND DELAY TRADEOFFS IN MANETS: A GLOBAL PERSPECTIVE

λ (i.i.d.)	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.

VIII. CONCLUSION

In this paper, we have studied the multicast capacity-delay tradeoffs in MANETs. Specifically, we established the upper bound on the optimal multicast capacity-delay tradeoffs under two-dimensional/one-dimensional i.i.d. 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 models constrain the direction of nodes' mobility, they achieve larger capacity than two-dimensional models. Our result is a vivid generalization of some early works in the area of unicast and multicast capacity, and will be further used to study hybrid random walk mobility models and the aggregate input capacity of heterogeneous networks in the future.

IX. Acknowledgment

This work is supported by National Fundamental Research Grant (2011CB302701, 2010CB731803, 2009CB3020402), NSF China (No. 60702046, 60832005, 60972050, 60632040); China Ministry of Education (No.20070248095); China Ministry of Education Fok Ying Tung Fund (No.122002); Qualcomm Research Grant; China International Science and Technology Cooperation Program (No. 2008DFA11630); PUJIANG Talents (08PJ14067); Shanghai Innovation Key Project (08511500400); National Key Project of China (2009ZX03003-006-03, 2009ZX03002-003, 2009ZX03002-005, 2010ZX03003-001-01); National High Tech Grant of China (2009AA01Z248, 2009AA1Z249, 2009AA011802); Shanghai Jiao Tong University Undergraduate Participation in Research Program.

References

[1] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," in *IEEE Transactions on Information 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 Transactions on Information 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. 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 trade-offs in mobile ad-hoc networks," in *IEEE Transactions on Information 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 Transactions on Information 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," IEEE International Conference on Computer Communications (INFOCOM'10), San Diego, CA, March 15-19, 2010.

[8] U. Kozat and L. Tassiulas, "Throughput Capacity of Random Ad Hoc Networks with Infrastructure Support." in *Proceedings of 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.

[10] W. Huang, X. Wang, Q. Zhang, "Capacity Scaling in Mobile Wireless Ad Hoc Network with Infrastructure Support," to appear in IEEE ICDCS 2010, Genoa, Italy, 2010.

[11] X. Li, S. Tang and O. Frieder, "Multicast capacity for large scale wireless ad hoc networks," in *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] X-Y. Li, Y. Liu, S. Li, and S. Tang, "Multicast Capacity of Wireless Ad Hoc Networks Under Gaussian Channel Model", IEEE/ACM Transactions on Networking (TON), Vol. 18, No. 4, August 2010, Pages 1145-1157.

[14] C. Hu, X. Wang and F. Wu, "MotionCast: On the Capacity and Delay Tradeoffs," in *ACM MobiHoc 09*, 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 Transactions on Mobile Computing*, 8(12):1585-1595, December 2009.

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

[18] C. Wang, X-Y. Li, C. Jiang, S. Tang, Y. Liu, "Scaling Laws on Multicast Capacity of Large Scale Wireless Networks", IEEE INFOCOM 2009, Rio de Janeiro, Brazil, April 2009.

[19] C. Hu, X. Wang, D. Nie, J. Zhao, "Multicast Scaling Laws with Hierarchical Cooperation," in Proc. of IEEE INFOCOM 2010, San Diego, USA, 2010.

[20] P. Li and Y. Fang, "The Capacity of Heterogeneous Wireless Networks," IEEE International Conference on Computer Communications (INFOCOM'10), San Diego, CA, March 15-19, 2010.

Appendix A – Proof of *Proposition 2*

We will focus on the case where 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 our scheme can obtain a capacity of $\frac{W}{32n^{(3s+2\alpha-2)/3}\log n}$ w.h.p. under *multicast* traffic pattern.

Lemma 5: 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 the expectation of B_i is $\mathbb{E}[B_i] = \frac{np}{m}$. And as $n \to \infty$, we have

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

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

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

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

(c) 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 $p = 1$.

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

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 5* (a), we have

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

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

Analysis of *Capture* is similar.

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