Multicast Capacity for Large Scale Wireless Ad Hoc Networks

Xiang-Yang Li
Department of Computer Science
Illinois Institute of Technology.
Chicago, IL, USA
xli@cs.iit.edu

***Abstract*—In this paper, we study the capacity of a large-scale random wireless network for multicast. Assume that n wireless nodes are randomly deployed in a square region with side-length a and all nodes have the uniform transmission range r and uniform interference range $R > r$. We further assume that each wireless node can transmit (or receive) at W bits/second over a *common* wireless channel. For each node v_i, we *randomly* and *independently* pick $k-1$ points $p_{i,j}$ $(1 \leq j \leq k-1)$ from the square, and then multicast data to the nearest node for each $p_{i,j}$. The aggregated multicast capacity is defined as the total data rate of all multicast sessions in the network. In this paper we derive matching asymptotic upper bounds and lower bounds on multicast capacity of random wireless networks. We show that the total multicast capacity is $\Theta(\sqrt{\frac{n}{\log n}} \cdot \frac{W}{\sqrt{k}})$ when $k = O(\frac{n}{\log n})$; [1] the total multicast capacity is $\Theta(W)$ when $k = \Omega(\frac{n}{\log n})$. Our bounds unify the previous capacity bounds on unicast (when $k = 2$) by Gupta and Kumar [6] and the capacity bounds on broadcast (when $k = n$) in [9], [18]. We also study the capacity of group-multicast for wireless networks where for each source node, we randomly select $k-1$ groups of nodes as receivers and the nodes in each group are within a constant hops from the group leader. The same asymptotic upper bounds and lower bounds still hold. For arbitrary networks, we provide a constructive lower bound $\Omega(\frac{\sqrt{n}}{\sqrt{k}} \cdot W)$ for aggregated multicast capacity when we can carefully place nodes and schedule node transmissions.**

***Index Terms*—Wireless ad hoc networks, capacity, multicast, broadcast, unicast, scheduling, optimization, VC dimension.**

I. Introduction

In wireless ad hoc networks, wireless nodes may cooperate in routing each others' packets. Lack of a centralized control of the functionality and possible node mobility give rise to many challenging issues at the network layer, the medium access layer, and physical layer of a wireless ad hoc network. At the network layer, a main challenging problem is that of routing, which has to deal with time-varying network topology, possible power-constraints of wireless nodes, and the unique characteristics of the wireless channel (such as unstable, broadcast nature, fading and so on). The choice of medium access control is also restricted by the fact that the network topology is time-varying, and there is no centralized control. In the literature, a number of results have been proposed to use the TDMA, CDMA, FDMA, and the dynamic assignment of frequency bands to improve the network throughput. Notice that for a mobile wireless network, a random medium access control protocol appears to be a favorite due to its simplicity and quick adaption to mobility and dynamic data rate by nodes. For a mobile wireless network, static FDMA is inefficient in dense networks, CDMA is very difficult to implement due to node mobility and the need for keeping track of spreading codes for nodes in the time-varying neighborhoods. Notice that TDMA has recently been proposed to improve the network throughput for some networks or partial of the networks [1], especially for static networks. At the physical layer an important issue is the power-control, which has been studied extensively in the literature. A careful selection of the transmission power of nodes can not only improve the nodal life, but also improve the spatial reuse of frequency and consequently possibly improve the network throughput.

In many applications, *e.g.*, wireless sensor networks, we often need a rough estimation on the achievable throughput when we randomly deploy n wireless nodes in a given region. The main purpose of this paper is to study the *asymptotic capacity* of large scale random wireless networks when we choose the *best* protocols for all layers. As in the literature, we will mainly consider one type of networks, large scale *random networks*, where a large number of nodes are randomly placed in the deployment region. We will study the capacity of a given wireless network where the nodes positions are randomly given a priori, and how the capacity of wireless networks scale with the number of nodes in the networks (when given a fixed deployment region), or scale with the size of the deployment region (when given a fixed deployment density) for multicast, which is a generalization of various number of operations such as unicast and broadcast. We assume that a set of n wireless nodes $V = \{v_1, v_2, \cdots, v_n\}$ are *randomly* distributed (with uniform distribution) in a square region with a side-length a and all nodes have the same transmission range r. For majority results presented in this paper, we assume that the deployment region a and the transmission range r are selected such that the resulted network will be connected with high probability[2] (*w.h.p.*). The results derived under this model also imply the

[1] A function $f(n) = O(g(n))$ if there exist $N_0 > 0$ and $c > 0$ such that for all $n > N_0$, $f(n) \leq c \cdot g(n)$. A function $f(n) = \Omega(g(n))$ if there exist $N_0 > 0$ and $c > 0$ such that for all $n > N_0$, $f(n) \geq c \cdot g(n)$. A function $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$.

[2] Here an event is said to happen with high probability, if for any $0 < \epsilon < 1$, there is a large integer N such that for a random network of size at least N, the probability that the event happens is at least $1 - \epsilon$.

same results for the dense model, when n nodes are distributed in a fixed region (such as a unit square by a proper scaling) and the uniform transmission range of all nodes are selected as the critical transmission range (CTR) to get a connected network with high probability.

In this paper, we will concentrate on the *multicast capacity* of a random wireless network, which generalizes both the unicast capacity [6] and broadcast capacity [9], [18] for random networks. Assume that a subset $\mathcal{S} \subseteq V$ of $n_s = |\mathcal{S}|$ nodes will serve as the source nodes of n_s multicast sessions. Each node $v_i \in \mathcal{S}$ *randomly and independently* chooses $n_d = k-1$ points $P_i = \{p_{i,j} \mid 1 \le j \le k-1\}$ in the square. For each point $p_{i,j}$, let $v_{i,j}$ be the node from V that is the closest to $p_{i,j}$. Then node v_i will send data to these $k-1$ nodes $U_i = \{v_{i,j} \mid 1 \le j \le k-1\}$ at an arbitrary data rate λ_i. The aggregated multicast capacity with $\mathcal{S}$ as roots for a network is defined as $\Lambda_{k,\mathcal{S}}(n) = \sum_{v_i \in \mathcal{S}} \lambda_i$ when there is a schedule of transmissions such that all multicast flows will be received by their destination nodes successfully within a finite delay.

Due to spatial separation, several wireless nodes can transmit simultaneously provided that these transmissions will not cause *destructive* wireless interferences to any of the transmissions. To describe when a transmission is received successfully by its intended recipient, we will allow one possible model for a successful one-hop reception: *protocol model*. We assume that each node $v \in V$ has a fixed constant transmission range r and a fixed constant interference range $R > r$. A node u can successfully receive a transmission from another node v iff there is *no* other node w such that $\|w-u\| \le R$ and node w is transmitting simultaneously with node v. Here $\|w-u\|$ is the Euclidean distance between w and u.

We assume the following simple wireless channel model as in the literature: each wireless node can transmit at W bits/second over a common wireless channel. For presentation simplicity, we assume that there is only one channel in the wireless networks. As always, we assume that the packets are sent from node to node in a multi-hop manner until they reach their final destinations. The packets could be buffered at intermediate nodes while awaiting for transmission. In this paper, we assume that the buffer is large enough so packets will not get dropped by any intermediate node. We leave it as future work to study the scenario when the buffers of intermediate nodes are bounded by some values. In some results, we assume that every intermediate node have infinite buffer size. For most of the results presented here, the delay of the routing is not considered, *i.e.*, the delay in the worst case could be arbitrarily large for some results.

Our Main Contributions: In this paper we propose two regimes for multicast capacity in terms of k. We derive matching analytical upper bounds and lower bounds on multicast capacity of a random wireless network. Assume that the side-length a of the deployment square and the transmission range r are selected [3] such that the network is connected with high probability, *i.e.*, $\frac{a}{r} = \Theta(\sqrt{\frac{n}{\log n}})$. We show that the aggregated multicast capacity of n multicast sessions is

$$\Lambda_k(n) = \begin{cases} \Theta(\sqrt{\frac{n}{\log n}} \cdot \frac{W}{\sqrt{k}}) & \text{when } k = O(\frac{n}{\log n}), \\ \Theta(W) & \text{when } k = \Omega(\frac{n}{\log n}) \end{cases} \quad (1)$$

Our bounds unify the previous capacity bounds on unicast (when $k = 2$) by Gupta and Kumar [6] and the capacity bounds on broadcast (when $k = n$) in [9], [18]. More generally, we prove that the aggregated multicast capacity of n_s random multicast sessions has the *same* asymptotic upper-bound as formula (1), and has the same asymptotic lower-bound as formula (1) whenever, $n_s \ge \Theta(\log k \cdot \sqrt{\frac{n \log n}{k}})$. Consequently, the per-node multicast capacity $\lambda_k(n)$ of n multicast sessions (with $k-1$ receivers per multicast session) is

$$\lambda_k(n) = \begin{cases} \Theta(\sqrt{\frac{1}{n \log n}} \cdot \frac{W}{\sqrt{k}}) & \text{when } k = O(\frac{n}{\log n}), \\ \Theta(\frac{W}{n}) & \text{when } k = \Omega(\frac{n}{\log n}) \end{cases} \quad (2)$$

The above capacity bounds are implied by a more general result for the following network setting: there are n_s multicast sessions, each with $k-1$ receivers from V, and the transmission range r and side-length a of the deployment square satisfying that the resulted random network is connected with high probability. Generally, when $n_s \ge \Omega(\log k \cdot \sqrt{\frac{n \log n}{k}})$, we prove that the aggregated multicast capacity of n_s multicast sessions is

$$\Lambda_k(n) = \begin{cases} \Theta(\frac{a}{r} \cdot \frac{W}{\sqrt{k}}) & \text{when } k = O(\frac{a^2}{r^2}) \\ \Theta(W) & \text{when } k = \Omega(\frac{a^2}{r^2}) \end{cases} \quad (3)$$

We also study the multicast capacity for group-multicast where, for each source node, we randomly select $k-1$ groups of nodes as receivers and the nodes in each group are within a constant number of hops from the group leader (who are nodes closest to $k-1$ randomly chosen points). We show that the asymptotic multicast capacity is still $\Theta(\sqrt{\frac{n}{\log n}} \cdot \frac{W}{\sqrt{k}})$ when $k = O(\frac{n}{\log n})$; and is $\Theta(W)$ when $k = \Omega(\frac{n}{\log n})$. For multicast in arbitrary networks, we provide a constructive lower bound $\Omega(\frac{\sqrt{n}}{\sqrt{k}} \cdot W)$ when we can carefully place nodes and schedule node transmissions.

The rest of the paper is organized as follows. In Section II we discuss in detail the network model and the channel model used in this paper. We briefly overview the proof techniques used to analyze the capacity upper-bound in Section III. In Section IV, we first present some upper-bounds on multicast capacity for random networks. In Section V, we then present an efficient method for multicast and prove that the capacity

[3]There are two scenarios here. The first case is that the deployment region is fixed as a unit square while the transmission range r of each wireless node is adjusted as $\Theta(\sqrt{\frac{\log n}{n}})$ such that the random network is connected with high probability. The second case is that the transmission range r is fixed as one unit, while the side-length a of the deployment square is adjusted to its critical value $\Theta(\sqrt{\frac{n}{\log n}})$ such that the random network is connected with high probability.

achieved by this method asymptotically matches the upper-bounds derived before. In Section VI, we study the multicast capacity bounds for group-multicast and the multicast capacity bounds for arbitrarily networks. We review the related results on network capacities in Section VII and conclude the paper in Section VIII with the discussion of some possible future works.

II. Network Model

The capacity of random wireless networks was first studied in a landmark seminar work by Gupta and Kumar [6]. There are different approaches to increase the network throughput, such as reducing the interference, the scheduling on the MAC layer, route selection on the routing layer, channel assignment if multi-channels are available, and power control on the physical layer. In this section, we first introduce our network system model, then we discuss in detail the interference models we will use and then define the problem that we will study in this paper.

We consider large scale random networks. Typically there are three ways to increase the number of network nodes to infinity.

1) One is to *fix* the deployment region and then increase the node density to infinity. This is typically called the *dense model*. This model is widely studied, *e.g.*, Gupta and Kumar studied the critical transmission range (CTR) [7] and the capacity for unicast [6] using this model. Compared with the practical deployment, this model has a drawback that the minimum power needed for having a connected network will be arbitrarily small when node density is sufficiently large.
2) Another way is to *fix* the node density to a given constant and increase the deployment region to infinity. This is typically called the *extended model*. Notice that to get a connected network with high probability, we also need to increase the transmission range of nodes. This model is also used by several papers to study the CTR or capacity, *e.g.*, [15], [22]. Compared with a practical deployment, this model also has a drawback that the minimum power needed for having a connected network will go to infinity when the area of the deployment region goes to infinity. Here we assume that there is a constant lower bound on the minimum SINR such that the receiver can correctly decode the signal.
3) The third way is to fix the transmission range of all nodes to some constant, then increase the node density (asymptotically same as the node degree when the transmission range is fixed) and the deployment area to increase the number of nodes in the network. We call this model the *constant-range model*. Assume that n nodes will be deployed. It has been proved in [21] that the minimum node degree for connectivity is $\Theta(\log n)$. This implies that the area of the deployment region is at most $\Theta(\frac{n}{\log n})$.

In this paper, we will adopt the third model. Notice that our results presented in this paper actually are immaterial to the model used. Most results presented in this paper rely on the ratio $\frac{a}{r}$ where a is the side-length of the deployment square and r is the transmission range, where either a or r or both could be a function of n.

In this paper, we assume that there is a set $V = \{v_1, v_2, \cdots, v_n\}$ of n communication terminals deployed in a region Ω. We mainly focus on the scenario when Ω is a square with side length a. Every wireless node has a transmission range r such that two nodes u and v can communicate directly if $\|u-v\| \leq r$ and there is no other interference. The complete communication graph is a undirected graph $G = (V, E)$, where E is the set of possible communication links uv with $\|u - v\| \leq r$. In this paper, we mainly assume that the transmission range r is a constant. Under this assumption, the side length a of the deployment square region Ω will be a function of n.

To schedule two links at the same time slot, we must ensure that the schedule will avoid interference. Several different interference models have been used to model the interferences in wireless networks. In this paper, we will mainly focus on the protocol interference model. We assume that each node v_i has a constant interference range R. Here any node v_j will be interfered by the signal from v_k if $\|v_k - v_j\| \leq R$ and node v_k is sending signal to some node other than v_j. In this paper, we always assume that the interference range R is within a small constant factor of the transmission range r, *i.e.*, $R \leq \beta \cdot r$ for a constant $\beta \geq 1$.

Capacity Definition: We assume that each node v_i could serve as the source node for some multicast. Assume that a subset $\mathcal{S} \subseteq V$ of $n_s = |\mathcal{S}|$ nodes will serve as the source nodes of n_s multicast sessions. For each node v_i, we *randomly and independently* choose a set P_i of $n_d = k - 1$ points $P_i = \{p_{i,j} \mid 1 \leq j \leq k-1\}$ in the deployment square. For each point $p_{i,j}$, let $v_{i,j}$ be the node from V that is the closest to $p_{i,j}$ (arbitrarily choose one if there are multiple nodes with the same shortest distance). Nodes $U_i = \{v_{i,j} \mid 1 \leq j \leq k-1\}$ will be the destination nodes of multicast from v_i. Then node v_i will send data to these $k-1$ nodes U_i at an arbitrary data rate λ_i. Notice that when the receivers are far away from the source node, we need multiple intermediate nodes to relay the data for v_i.

Given the set $\mathcal{S}$ of $n_s = |\mathcal{S}|$ source nodes, let $\lambda_{\mathcal{S}} = (\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_{n_s-1}}, \lambda_{n_s})$ be the *rate vector* of the multicast data rate of all n_s multicast sessions. Here λ_{i_j} is the data rate of node $v_{i_j} \in \mathcal{S}$, for $1 \leq j \leq n_s$. When given a *fixed* network $G = (V, E)$, where the node positions of all nodes V, the set of receivers U_i for each source node v_i, and the multicast data rate λ_i for each source node v_i are all fixed, we first define what is a feasible rate vector λ for the network G.

Definition 1 (Feasible Rate Vector): Given set $\mathcal{S}$ of n_s source nodes, a multicast rate vector $\lambda_{\mathcal{S}}$ bits/sec is *feasible* if there is a spatial and temporal scheme for scheduling transmissions such that by operating the network in a multi-hop fashion and buffering at intermediate nodes when awaiting transmission, every node v_i can send λ_i bits/sec average to its chosen $k-1$ destination nodes. That is, there is a $T < \infty$ such that in every time interval (with unit seconds) $[(i-1)\cdot T, i\cdot T]$, every node $v_i \in \mathcal{S}$ can send $T \cdot \lambda_i$ bits to its corresponding $k-1$ receivers U_i.

The total throughput capacity of such feasible rate vector for multicast is defined as

$$\Lambda_{k,\mathcal{S}}(n) = \sum_{v_i \in \mathcal{S}} \lambda_i. \tag{4}$$

The average per-flow multicast throughput capacity is

$$\lambda^a_{k,\mathcal{S}}(n) = \frac{\sum_{v_i \in \mathcal{S}} \lambda_i}{n_s}, \tag{5}$$

The minimum per-flow multicast throughput capacity is

$$\lambda_{k,\mathcal{S}}(n) = \min_{v_i \in \mathcal{S}} \lambda_i, \tag{6}$$

where k is the total number of nodes in each multicast session, including the source node. When $\mathcal{S}$ is clear from the context, we drop $\mathcal{S}$ from our notations. When we mention *per flow multicast capacity*, hereafter we mean the minimum per flow multicast capacity, if not explained otherwise.

Definition 2 (Throughput Capacity): An aggregated multicast throughput $\Lambda_k(n)$ bits/sec is *feasible* for n_s multicast sessions (each session with k terminals) if there is a rate vector $\lambda_{\mathcal{S}} = (\lambda_{i_1}, \lambda_{i_2}, \cdots, \lambda_{i_{n_s-1}}, \lambda_{n_s})$ that is feasible and $\Lambda_k(n) = \sum_{v_i \in \mathcal{S}} \lambda_i$. Similarly, we say $\lambda_k(n) = \min_{v_i \in \mathcal{S}} \lambda_i$ is a feasible per-flow multicast throughput capacity.

Definition 3 (Capacity of Random Networks): The *aggregated multicast capacity* of a class of random networks is of order $\Theta(g(n))$ bits/sec if there are deterministic constants $c > 0$ and $c < c' < +\infty$ such that

$$\begin{aligned} \lim_{n\to\infty} \mathbf{Pr}\left(\Lambda_k(n) = cg(n) \ is \ feasible\right) &= 1 \\ \liminf_{n\to\infty} \mathbf{Pr}\left(\Lambda_k(n) = c'g(n) \ is \ feasible\right) &< 1 \end{aligned}$$

We say that the *multicast capacity per flow* of a class of random networks is of order $\Theta(f(n))$ bits/sec if there are deterministic constants $c > 0$ and $c < c' < +\infty$ such that

$$\begin{aligned} \lim_{n\to\infty} \mathbf{Pr}\left(\lambda_k(n) = cf(n) \ is \ feasible\right) &= 1 \\ \liminf_{n\to\infty} \mathbf{Pr}\left(\lambda_k(n) = c'f(n) \ is \ feasible\right) &< 1 \end{aligned}$$

Here the probability is computed using *all* possible connected random networks formed by n nodes distributed in a square with side-length a.

Useful Known Results: Throughput this paper, we will repeatedly use the following results from probability theory literature.

Lemma 1 (Chebyshev's Inequality): For a variable X, and $A > 0$

$$\mathbf{Pr}\left(|X - \mu| \ge A\right) \le \frac{\mathrm{Var}(X)}{A^2},$$

where $\mu = E(X)$, $\mathrm{Var}(X)$ is the variance of X.

Lemma 2 (Weak Law of large numbers): Consider n uncorrelated variables X_i, $1 \le i \le n$ with same expected value $\mu = E(X_i)$ and variance $\sigma^2 = \mathrm{Var}(X_i)$. Let $\overline{X} = \frac{\sum_{i=1}^n X_i}{n}$. Then for any $\epsilon > 0$,

$$\mathbf{Pr}\left(|\overline{X} - \mu| < \epsilon\right) \ge 1 - \frac{\sigma^2}{n \cdot \epsilon^2}.$$

Lemma 3 (Hoeffding's inequality): Consider n independent variables X_i with $\mathbf{Pr}\left(X_i \in [a_i, b_i]\right) = 1$. Let $X = \sum_{i=1}^n X_i$. Then

$$\mathbf{Pr}\left(X - E(X) \ge nt\right) \le e^{-\frac{-2n^2t^2}{\sum_{i=1}^n (b_i - a_i)^2}}, \text{ when } 0 < t.$$

Lemma 4 (Binomial Distribution): Consider n independent variables $X_i \in \{0, 1\}$ with $p = \mathbf{Pr}\left(X_i = 1\right)$. Let $X = \sum_{i=1}^n X_i$. Then

$$\begin{aligned} \mathbf{Pr}\left(X \le \xi\right) &\le e^{\frac{-2(n\cdot p - \xi)^2}{n}}, \quad \text{when } 0 < \xi \le n \cdot p. \\ \mathbf{Pr}\left(X > \xi\right) &< \frac{\xi(1-p)}{(\xi - n\cdot p)^2}, \quad \text{when } \xi > n \cdot p. \end{aligned}$$

We will also use the uniform convergence in the weak law of large numbers. We recall the following definitions by Vapnik and Chervonenkis [19]. Let $\mathcal{U}$ be the input space. Let $\mathcal{C}$ be a family of subsets of $\mathcal{U}$. A *finite* set S (called sample in machine learning) is *shattered* by $\mathcal{C}$, if for *every* subset B of S, there *exists* a set $A \in \mathcal{C}$ such that $A \bigcap S = B$. The *VC-dimension* of $\mathcal{C}$, denoted by VC-d($\mathcal{C}$), is defined as the maximum value d such that there exists a set S with cardinality d that can be shattered by $\mathcal{C}$. For sets of *finite* VC-dimension, one has uniform convergence in the weak law of large numbers:

Theorem 5 (The Vapnik-Chervonenkis Theorem): If $\mathcal{C}$ is a set of *finite* VC-dimension VC-d($\mathcal{C}$), and $\{X_i \mid i = 1, 2\cdots, N\}$ is a sequence of *i.i.d.* random variables with *common* probability distribution P, then for every $\epsilon, \delta > 0$,

$$\mathbf{Pr}\left(\sup_{A\in\mathcal{C}} \left| \frac{\sum_{i=1}^N I(X_i \in A)}{N} - P(A) \right| \le \epsilon \right) > 1 - \delta \tag{7}$$

whenever

$$N > \max\left\{ \frac{8 \cdot \text{VC-d}(\mathcal{C})}{\epsilon} \cdot \log \frac{13}{\epsilon}, \frac{4}{\epsilon} \log \frac{2}{\delta} \right\}. \tag{8}$$

Here $I(X_i \in A)$ takes value 1 if $X_i \in A$ and 0 otherwise.

Notations: Throughput this paper, for a continuous region Ω, we use $|\Omega|$ to denote its area; for a discrete set S, we use $|S|$ to denote its cardinality; for a tree T, we use $\|T\|$ to denote its total Euclidean edge lengths; $x \to \infty$ denotes that variable x takes value to infinity.

III. General Techniques For Upper-Bounds

A. Simultaneous Transmissions

In previous studies [6], [16] of capacity of random networks, a common approach is to analyze the expected number of *hops* $H(b)$ a bit b has to travel and the total number of simultaneous transmissions $S = O(\frac{a^2}{r^2})$ possible in the system. If each source node generates data at rate λ, the number of bits generated by these n_s sources in time interval T is simply $\lambda T n_s$. Thus, the total number of transmissions of all bits to their destinations is $\lambda T n_s H(b)$ almost surely. Consequently, we have $\lambda T n_s H(b) \le T \cdot S$. This implies that $\lambda = O(\frac{a^2}{r^2} \cdot \frac{1}{n_s H(b)}) = O(\frac{n}{\log n} \cdot \frac{1}{n_s H(b)})$. In [6], for *unicast*, Gupta and Kumar essentially used $\Theta(\frac{1}{r})$ (assumed $a = 1$) as estimation of $H(b)$ and derived $\Theta(\frac{W}{n_s \cdot r})$ as per-node capacity upper-bound. In [16], for *multicast*, Shakkottai *et al.*

essentially used $H(b) = \Theta(\frac{\sqrt{k}}{r})$ (assumed $a = 1$) to derive $O(\frac{W}{n_s \cdot \sqrt{k} \cdot r}) = O(\frac{\sqrt{n}}{n_s \sqrt{k \log n}})$ as per-node capacity upper-bound. Although this traditional technique is valid and convenient for studying the asymptotic unicast capacity and the multicast capacity with some special configurations ($k = n^{1-\epsilon}$ for some $0 < \epsilon < 1$) [16], this may produce a pessimistic or even erroneous upper-bound for asymptotic multicast capacity in a general setting studied in this paper. For example, when $k = n$ (*i.e.*, broadcast), formula $O(\frac{\sqrt{n}}{n_s \sqrt{k \log n}})$ only produces an upper-bound $O(\frac{1}{n_s \sqrt{\log n}})$, which is asymptotically smaller than the achievable per-node broadcast capacity $\Theta(\frac{1}{n_s})$ implied in [9], [10], [18]. The reason for this discrepancy is that for a multicast tree T with total length $\|T\|$, value $\Theta(\frac{\|T\|}{r})$ may *not* give the lower bound on the number of transmissions needed by the tree T due to the multicast natural of wireless transmissions. For the example of broadcast, later we will show that for any broadcast tree T, $\|T\| = \Omega(\sqrt{n})$ almost surely. Thus, $\Theta(\frac{\|T\|}{r}) = \Omega(\frac{n}{\sqrt{\log n}})$ due to $r = \Theta(\sqrt{\frac{\log n}{n}})$. On the other hand, a simple broadcast based on a connected dominating set [9] will only require $\Theta(\frac{n}{\log n})$ transmissions.

To address the above challenges and discrepancies, we use two new approaches to analyze the upper-bound of multicast capacity:

1) **Area Argument**: This is based on the area covered by the transmission disks of all internal nodes in a multicast tree;
2) **Data Copies Argument**: This approach is based on the number of nodes that receive a copy of a multicast data during the transmissions of all nodes in the tree.

B. Shaded Area of Transmissions

The area argument essentially works as follows. When we multicast from one source node v_i to all its $k-1$ receivers U_i, all nodes lying inside the interference region of *any* transmitting node for this multicast session cannot receive data from other nodes simultaneously. We call the region where no node can receive data from other transmitting node when v_i is transmitting as *shaded* by the transmission of v_i. For any node u, let $t_i(u)$ be the time-intervals that node u will transmit data for multicast tree T_i. Thus, a multicast tree will claim a number of cylinders ($D(u, r) \times t_i(u)$ for internal node u in T) in the space-time dimension $\mathcal{R}^2 \times \mathcal{T}$, where $D(u, r)$ denotes the transmission disk of node u, $\mathcal{T}$ is the scheduling period. Thus, given a multicast tree T_i for multicast originated from v_i, the pairs of $(D(u, r), t_i(u))$ (*i.e.*, transmission disk $D(u, r)$ will be used for multicast originated at v_i during the transmission time-interval $t_i(u)$) claimed by this multicast should be disjoint from the pairs claimed by other multicast sessions.

Let A_i be the area of the overall region that is shaded by any transmitting node for the operation (unicast, broadcast, multicast, or any other operation) of node v_i, and λ_i be the data rate. Then obviously $\sum_{i=1}^{n} \lambda_i \cdot A_i \leq W \cdot \Phi$ where Φ is the total area of the region covered by the transmitting disks of all nodes. Thus, it is not difficult to prove the following lemma:

Lemma 6: For any operation $\mathcal{O}$, such as multicast, let A_i be the area of the region defined by uniting the transmission regions of all transmitting nodes for operation initiated by node v_i. If A_i is at least $\wp$ with high probability for every node v_i in $\mathcal{S}$, then, *w.h.p.*, the aggregated capacity for this operation in a random network deployed in a region with area Φ is at most $\frac{\Phi \cdot W}{\wp}$.

C. Amortized Receiving

The data-copies argument works as follows. When we multicast from one source node v_i to all its $k-1$ receivers U_i, it is more likely that other nodes will also get a copy of the data. Here, for the purpose of analysis, when a node v sends data to one of its neighboring nodes, **all** its neighboring nodes will be *charged* a *copy* of the data. Notice that here a neighboring node w may not be the intended receiver. However, since when v is transmitting, any of its neighboring node w **cannot** receive data simultaneously from any other transmitting node due to interference, we will say that node w also gets a copy of the data. For multicast with $k-1$ receivers, clearly, at least k nodes will get a copy of the data. Generally, assume that C_i nodes will get a copy of the data when the $k-1$ receivers are randomly selected for each possible source node v_i. Obviously, $\sum_{v_i \in \mathcal{S}} \lambda_i \cdot C_i \leq n \cdot W$. Further assume that $C_i \geq C$ almost surely, *i.e.*, $\mathbf{Pr}(C_i \geq C) \to 1$ as n or k goes to infinity. Then the total multicast capacity satisfies, almost surely,

$$\Lambda_k(n) = \sum_{v_i \in \mathcal{S}} \lambda_i \leq \frac{n \cdot W}{C}. \tag{9}$$

Clearly, $C \geq k$. Next subsection is devoted to give a better lower bound on C. The following lemma is straightforward.

Lemma 7: For any operation $\mathcal{O}$, such as multicast, let $\mathcal{X}$ be the number of nodes that will *receive* a *copy* of the data (*i.e.*, fall inside the interference region of any one of its transmitting nodes). Assume that $\mathcal{X}$ is at least $\mathcal{N}$ with high probability. Then, with high probability, the aggregated capacity for this operation by all nodes in a random network of n nodes is at most $\frac{n \cdot W}{\mathcal{N}}$.

In our proofs, we will utilize these two technical lemmas to give upper-bound on the capacity of a random network for an operation that will be performed by each node of the network, such as multicast. Notice that the above lemmas require us to find largest $\wp$ such that $\mathbf{Pr}(\mathcal{A} \geq \wp) \to 1$, or the largest $\mathcal{N}$ such that $\mathbf{Pr}(\mathcal{X} \geq \mathcal{N}) \to 1$. In some cases, such $\wp$ may be much smaller than the mean value $E(\mathcal{A})$ of $\mathcal{A}$; such $\mathcal{N}$ may be much smaller than the mean value $E(\mathcal{X})$ of $\mathcal{X}$. In these cases, we could rely on a much stronger technical lemmas based on *law of large numbers* when the number n_s of operations needed to perform goes to infinity.

IV. Upper Bounds on Multicast Capacity for Random Networks

A. The upper-bound on $\frac{a}{r}$

We assume that n wireless nodes V with transmission range r are randomly and uniformly distributed in a square region with side length a. We first study the asymptotic bound on a/r such that the resulted network $G = (V, E)$ is connected with probability going to 1 as n goes to infinity. Notice that for a set of nodes, the CTR for connectivity is always the length of the longest edge of the Euclidean minimum spanning tree (EMST) of this set of nodes [7], [14], [15]. Consequently, studying the CTR for connectivity is equivalent to studying the longest edge of the EMST of a set V of nodes when V follows a certain distribution such as Poisson distribution or random uniform distribution.

Assume n points are distributed uniformly at random in the 2-dimensional *unit square* and let M_n be the random variable denoting the length of the longest edge of EMST built on this set of n nodes. It was proved in [14] that, for any real number β,

$$\lim_{n\to\infty} \mathbf{Pr}\left(n\pi \cdot M_n^2 - \log n \le \beta\right) = \frac{1}{e^{e^{-\beta}}}.$$

Assume n points are distributed uniformly at random in the 2-dimensional square with side length a. Let $M_{n,a}$ be the random variable denoting the length of the longest edge of EMST built on this set of n nodes. Then a simple scaling shows that

$$\lim_{n\to\infty} \mathbf{Pr}\left(n\pi \cdot (\frac{M_{n,a}}{a})^2 - \log n \le \beta\right) = \frac{1}{e^{e^{-\beta}}}.$$

for any real number β. Thus, with probability $\frac{1}{e^{e^{-\beta}}}$, we know that the longest edge length $M_{n,a}$, of EMST built on n points distributed in a square with side-length a, is at most $\sqrt{\frac{\log n+\beta}{n\pi}} \cdot a$. Thus, when $\beta \to \infty$ and $a \le \sqrt{\frac{n\pi}{\log n+\beta}} \cdot r$, we know that the longest edge of EMST has length at most r *almost surely*. Thus, we have

Theorem 8: Assume that n nodes, each with transmission range r, are randomly uniformly deployed in a square region of side length a. When $\frac{a}{r} \le \sqrt{\frac{n\pi}{\log n+\beta}}$ for $\beta \to \infty$, the resulted network $G = (V, E)$ is connected with probability at least $\frac{1}{e^{e^{-\beta}}}$.

For example, we can set $a = r\sqrt{\frac{n}{\log n}}$ where $\beta = (\pi - 1)\log n$.

To our surprise, we find that the multicast capacity of a random network where each multicast session has $k-1$ receivers has two regimes: when the number of receivers $k-1$ is over some threshold, multicast capacity is asymptotically same as the broadcast capacity; otherwise, the multicast capacity decreases linearly over $\frac{1}{\sqrt{k}}$. In the next subsections, we will provide upper-bounds for each case separately.

B. When $k = O(a^2/r^2)$

We first study the multicast capacity when the number of receivers is at most $O(a^2/r^2)$. We will present upper bound of the total multicast capacity. A trivial upper bound for total multicast capacity is $W \cdot n$ since there are n source nodes and each source node can only send W bits/sec. A refined upper bound is $\Lambda_k(n) \le \frac{n \cdot W}{k}$ which is derived from the perspective of recipients: (1) each node can receive at most W bits/sec, and (2) among received data by all nodes, any data from any source node will have at least k copies (one copy at each of the $k-1$ receivers and one copy at the source node). From Lemma 7, for a multicast tree T_i spanning source node v_i and $k-1$ receivers U_i, we would like to know a lower bound on the number of internal nodes used in T_i. To analyze this value, we first study the asymptotic lower bound of the Euclidean length $\|T_i\|$ of a multicast tree T_i.

Lemma 9: [4] Given any k nodes U, *any* multicast tree (also called Steiner tree) spanning these k nodes (may be using some additional relay nodes) will have an Euclidean length at least $\varrho \cdot \|EMST(U)\|$, where $\varrho \ge \frac{\sqrt{3}}{2}$ and $EMST(U)$ is the EMST spanning U.

Observe that the tight bound on $\varrho = \frac{\sqrt{3}}{2}$ is the famous Gilbert and Pollak conjecture which was proved by Du and Hwang in 1992 [4]. A bound $\varrho \ge \frac{1}{2}$ can be easily proved as follows. For any steiner multicast tree T spanning these k nodes, we construct an Euler tour on this tree. Clearly the total length of the Euler tour (EC) is 2 times of the length of the multicast tree T. On the other hand, the Euler tour has length at least that of the Euclidean minimum spanning tree for these k nodes. The statement follows from $\|EMST\| < \|EC\| = 2 \cdot \|T\|$. Recall that in this paper, $\|T\|$ denotes the total Euclidean length of all links in a structure T.

Based on Lemma 9, to get a lower bound on $\|T_i\|$ of any multicast tree T_i, we need study the length of EMST spanning these k random nodes. In [17], Steele established the following result:

Lemma 10: The total edge length of the EMST of n nodes randomly and uniformly distributed in a d-dimensional cube of side-length a is asymptotic to $\tau(d) \cdot n^{\frac{d-1}{d}} \cdot a$, where $\tau(d)$ is a constant depending only on the dimension d.

Thus, based on Lemma 9 and Lemma 10, we have

Lemma 11: The total edge length, denoted by $\|T_i\|$, of *any* multicast tree T_i spanning k nodes that are randomly placed in a square of side-length a almost surely is at least, when $k \to \infty$,

$$\varrho \cdot \tau(2) \cdot \sqrt{k} \cdot a.$$

From now on, for simplicity, we will denote $\tau \leftarrow \sqrt{3}\tau(2)/2$. Thus, the total link length of a multicast tree is at least, almost surely, $\tau \cdot \sqrt{k} \cdot a$ when k goes to infinity.

Let $X = \|EMST(U)\|$ of a set of k randomly selected nodes U in a square of side-length a. It was shown in [17] that $\mathrm{Var}(X) \ll a^2 \cdot \log k$. We then show that $X \le 2\sqrt{2}\sqrt{k}a$.

Lemma 12: For any k nodes U placed in a square region with side-length a, the length of EMST spanning U is at most $2\sqrt{2}\sqrt{k}a$.

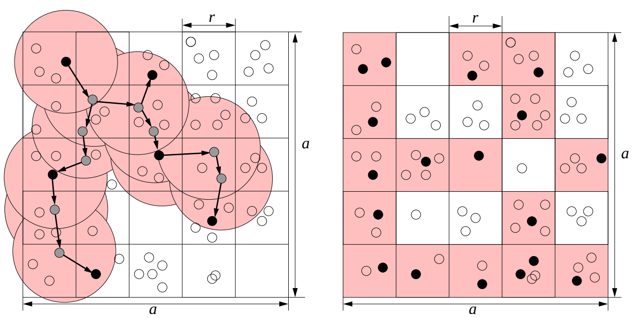
Proof: Given k nodes in the square, we will use Prim's algorithm to construct EMST: originally each node is a component, and then we iteratively find a shortest edge to connect two components to form a larger component until only one component is left. Consider the $(k+1-g)$-th step (for $1 \le g \le k$), which has j connected components as input. For $g \ge 2$, if we partition the square into a $\lfloor\sqrt{g-1}\rfloor$ by $\lfloor\sqrt{g-1}\rfloor$ grid with side-length $\frac{a}{\lfloor\sqrt{g-1}\rfloor}$, then there is at least one cell that contains at least two connected components. This implies that the shortest edge connecting components at the $(k+1-g)$th step is at most $\sqrt{2}\frac{a}{\lfloor\sqrt{g-1}\rfloor}$. Consequently, the EMST has length at most

$$\sum_{j=2}^{k} \sqrt{2}\frac{a}{\lfloor\sqrt{g-1}\rfloor} \quad \le \quad \sum_{i=1}^{\sqrt{k-1}} \sqrt{2}\frac{a\cdot((i+1)^2 - i^2 - 1)}{i}$$

$$\leq \quad 2\sqrt{2}\sqrt{k-1} \cdot a$$

This finishes the proof. ∎

Bound the data copies: A straightforward lower-bound on the number of nodes (including leaf nodes) needed in a multicast tree spanning k nodes randomly selected in a square of side-length a is $\tau \cdot \sqrt{k} \cdot \frac{a}{r}$ with high probability. This bound can be derived as follows: (1) the Euclidean length of a multicast tree is at least $\tau \cdot \sqrt{k} \cdot a$ with high probability, and (2) the transmission range of each node is only r, thus, removing one tree edge incident on a leaf node will reduce the total edge length by at most r and we will reduce the number of nodes by 1. Consequently, we have $C \geq \tau \cdot \sqrt{k} \cdot a/r$, with high probability. Although this bound on C is much better than bound $C \geq k$ when $k = O(a^2/r^2)$, the bound can be further improved based on the following observation. When nodes on the multicast tree relay data from the source node to receivers, not only its downstream nodes of the multicast tree will receive the data, but also all its neighboring nodes (in communication graph G) will get a *copy* of the data. We will then analyze the number of nodes that will get the copy of the data. Given a multicast tree T, let $D(T)$ be the region covered by all transmitting disks of all transmitting nodes (internal nodes of T) in the multicast tree T. Observe that the leaf nodes do *not* contribute to $D(T)$ at all here. See Figure 1 (a) for illustration. Clearly, the area of $D(T)$, denoted by $|D(T)|$,



(a) Region $D(T)$ (b) Dense receivers

Fig. 1. (a) Region $D(T)$ covered by transmitting disks of internal nodes in multicast tree T. Here the solid black nodes are receivers/source and gray nodes are Steiner nodes. (b) Partition of square with side-length a into squarelets with side-length r. Here the solid black nodes are receivers/source. Shaded squarelets are squarelets with at least one receiver.

is at most $|D(T)| \leq 2r \cdot \|T\| + k \cdot \pi r^2/2$ where $\|T\|$ is the total Euclidean length of all links in T. We will prove that the area of $D(T)$ is also at least $\frac{\tau\sqrt{k}a \cdot r}{c_0}$, *w.h.p.*, for some constant c_0 independent of the network.

Lemma 13: The area of the region $D(T)$, denoted by $|D(T)|$, with high probability, is at least $\frac{\tau\sqrt{k}a \cdot r}{c_0}$ when the number of receivers/source nodes $k < (\frac{\tau(1-(6(d+1)\cdot\rho))}{6(d+1)+1})^2 \cdot \frac{a^2}{r^2}$, for some constant $c_0 = 1/(\rho\pi)$, where $0 < \rho < \frac{1}{6(d+1)}$ and constant $d \leq 13$.

Proof: For any multicast tree T spanning source node v_i and the set of receivers U_i, for convenience, let $V(T)$ be the set of nodes in tree T; let $U_i' = U_i \cup \{v_i\}$; let $I(T)$ be all the Steiner nodes used to connect them, *i.e.*, $I(T) = V(T) \backslash U_i'$. Clearly the communication graph defined on $V(T)$ (where two nodes are connected iff their Euclidean distance is no more than r) is connected. We use G_T to denote such induced graph. We will then build another multicast tree T' from G_T to connect nodes U_i'.

In graph G_T, we build a connected dominating set (CDS) using a method described in [3], [20]. Source node s will be added to the CDS if it is not in the CDS. It has been proved in [3], [20] that, in the constructed CDS, each node on the CDS has a degree bounded by a constant, say d. For example, it can be shown that the degree of node in CDS is bounded by 13 if the method presented in [20] is used. The multicast tree T' is then a simple breadth-first-search tree computed from the CDS, rooted at the source node s.

We essentially will prove that each point from the region $D(T')$ is covered by at most a constant c_0 number of disks from the multicast tree T'. For each point p in the region, we divide the disk $D(p, r)$ centered at point p with radius r into 6 equal sized sectors. Thus, any pair of nodes falling into the same sector will be within distance of r of each other, and thus connected in the original communication graph. Consequently, for each point p, the number of disks from $D(T')$ that cover p is at most $6(d+1)$. If it is at least $6(d+1)+1$, then at least one of the sectors will have at least $d+2$ nodes, which implies that any node in that sector will have degree at least $d+1$ in the induced CDS graph. This is a contradiction to the fact that the degree of induced CDS is bounded by d. Thus, the area of the region $D(T')$ is thus at least $\frac{|I(T')| \cdot \pi r^2}{6(d+1)}$, where d is the degree bound on the induced CDS graph constructed. Here $|I(T')|$ is the number of internal nodes in multicast tree T'.

Notice that some leaf nodes in T may become internal nodes in T'; some internal nodes may not be used by tree T' at all. Let $A(T)$ be the region covered by all disks centered at all nodes of a tree T, including the leaf nodes. Let $\ell(T)$ be the number of leaf nodes in a tree T. Obviously, $|D(T)| + \ell(T) \cdot \pi r^2 \geq |A(T)| \geq |A(T')| \geq |D(T')|$. Thus,

$$|D(T)| \geq |D(T')| - \ell(T) \cdot \pi r^2 \geq \frac{|I(T')| \cdot \pi r^2}{6(d+1)} - \ell(T) \cdot \pi r^2.$$

Obviously, $\ell(T) \leq k$. For a multicast tree T', there are at most k leaf nodes. If we remove all edges in T' incident on leaf nodes, the total edge length of all edges left is at least $|T'| - k \cdot r$. Thus, the number of internal nodes $|I(T')|$ in T' is at least $\frac{|T'| - k \cdot r}{r}$. Notice that T' is a tree spanning the source node s and all receivers U_i'. Thus, with high probability, $|T'| \geq \tau\sqrt{k} \cdot a$ since U_i' has k nodes. Thus, with high probability, we have

$$|I(T')| \geq \frac{\tau\sqrt{k} \cdot a}{r} - k.$$

Assume that $k < (\frac{\tau(1-(6(d+1)\cdot\rho))}{6(d+1)+1})^2 \cdot \frac{a^2}{r^2}$, which implies

$$\begin{aligned} |D(T)| &\geq \frac{|I(T')| \cdot \pi r^2}{6(d+1)} - \ell(T) \cdot \pi r^2 \\ &\geq \left(\frac{\frac{\tau\sqrt{k} \cdot a}{r} - k}{6(d+1)} - k \right) \cdot \pi r^2 \\ &\geq \rho\pi\tau\sqrt{k} \cdot a \cdot r \end{aligned}$$

For example, we can set $\rho = \frac{1}{12(d+1)}$. This finishes the proof. ■

For convenience, hereafter, we use

$$\theta_1 = \left(\frac{\tau(1-(6(d+1)\cdot\rho))}{6(d+1)+1}\right)^2$$

to denote the threshold value such that Lemma 13 is true if $k < \theta_1 \cdot \frac{a^2}{r^2}$. Based on Lemma 13, we know that the expected number, denoted by C, of nodes from V that is in the region $D(T)$ is at least

$$|D(T)| \cdot \frac{n}{a^2} \geq \frac{\tau \cdot \sqrt{k} \cdot a \cdot r \cdot n}{c_0 a^2} = \frac{\tau \cdot \sqrt{k} \cdot r \cdot n}{c_0 a}$$

Recall that we assume that there is only one single channel in the network. It is then not difficult to show the following lemma:

Lemma 14: With high probability, the number C of nodes that get a copy of the multicast data satisfies $C > \frac{\tau \cdot r \cdot \sqrt{k} \cdot n}{2c_0 a}$.

Proof: Consider a multicast tree T. Notice that n wireless nodes will be randomly distributed in a square region of side-length a. Let $X_i = \{0, 1\}$ be an indicator variable whether the ith node v_i will fall inside the region $D(T)$ for a multicast tree T. Clearly $\mathbf{Pr}(X_i = 1) = \frac{|D(T)|}{a^2}$. Recall that, we already proved that, with high probability, $|D(T)| \geq \frac{\tau\sqrt{k}a \cdot r}{c_0}$. Thus, we have

$$\mathbf{Pr}(X_i = 1) \geq \frac{\tau\sqrt{k} \cdot r}{c_0 \cdot a}.$$

Obviously, $X = \sum_{i=1}^{n} X_i$ is the expected number of nodes falling inside the region $D(T)$, which is also the number C of nodes that will get a copy of the data by multicast. Then the expected value $E(X) \geq \frac{\tau\sqrt{k} \cdot r \cdot n}{c_0 \cdot a}$.

Based on Lemma 4, we have

$$\mathbf{Pr}\left(C \leq n \cdot \frac{|D(T)|}{2a^2}\right) \leq e^{\frac{-2\left(n \cdot \frac{|D(T)|}{a^2} - n \cdot \frac{|D(T)|}{2a^2}\right)^2}{n}}$$
$$= e^{\frac{-n \cdot |D(T)|^2}{2a^4}} \leq e^{-\frac{n \cdot \tau^2 \cdot k \cdot r^2}{2(c_0)^2 \cdot a^2}}$$

Notice that to guarantee a connected network with high probability, we have $a < \sqrt{\frac{\pi n}{\log n}} \cdot r$ with high probability. Thus, when $n \to \infty$ we have

$$\mathbf{Pr}\left(C \leq n \cdot \frac{|D(T)|}{2a^2}\right) \leq e^{-\frac{n \cdot \tau^2 \cdot k \cdot r^2}{2(c_0)^2 \cdot a^2}} \leq e^{-\frac{\tau^2 \cdot k \cdot \log n}{2\pi c_0^2}} = \frac{1}{n^{\frac{\tau^2 \cdot k}{2\pi c_0^2}}}$$

Consequently, when $n \to \infty$, $\mathbf{Pr}\left(C \leq n \cdot \frac{|D(T)|}{2a^2}\right) \to 0$. Thus,

$$\mathbf{Pr}\left(C > \frac{\tau \cdot r \cdot \sqrt{k} \cdot n}{2c_0 \cdot a}\right) \geq \mathbf{Pr}\left(C > n \cdot \frac{|D(T)|}{2a^2}\right) \to 1.$$

This finishes the proof. ■

Consequently, we have the following theorem:

Theorem 15: The multicast capacity with $k-1$ receivers for n nodes that are randomly and uniformly deployed in a square with side-length a is at most $c_1 \cdot \frac{aW}{r\sqrt{k}}$ for some constant c_1 when $k < \theta_1 \cdot a^2/r^2$.

Proof: Notice that the multicast capacity is at most $\frac{nW}{C}$ and, with high probability, $C \geq \frac{\tau \cdot r \cdot \sqrt{k} \cdot n}{2c_0 a}$ when $k < \theta_1 \cdot a^2/r^2$. Thus, the multicast capacity $\Lambda_k(n)$ is at most

$$\frac{nW \cdot 2c_0 a}{\tau \cdot r \cdot \sqrt{k} \cdot n} = c_1 \cdot \frac{aW}{r\sqrt{k}}$$

for a constant $c_1 = \frac{2c_0}{\tau}$. This finishes the proof. ■

Recall that we have proved that, to guarantee that we have a connected network with high probability, we need $a \leq r\sqrt{\frac{n\pi}{\log n + \beta}}$ for $\beta \to \infty$. Thus, letting $c_2 = c_1\sqrt{\pi}$, we have the following theory:

Theorem 16: The multicast capacity for a random network of n nodes, when $k < \theta_1 \cdot a^2/r^2$, is at most

$$\Lambda_k(n) \leq c_2 \cdot \frac{\sqrt{n}}{\sqrt{\log n} \cdot \sqrt{k}} \cdot W = O(\frac{\sqrt{n}}{\sqrt{\log n} \cdot \sqrt{k}} \cdot W).$$

With n_s multicast sessions, the per flow multicast capacity is at most

$$\lambda_k(n) = \min\{W, \frac{\Lambda_k(n)}{n_s}\} = \min\{W, O(\frac{\sqrt{n}}{n_s\sqrt{\log n} \cdot \sqrt{k}} \cdot W)\}.$$

Notice that Theorem 15 was proved under the assumption that $k \to \infty$. When this is not the case, we can prove that the per-flow multicast capacity λ (when each source node generates multicast data at rate λ) also satisfies that $\lambda_k(n) = O(\frac{\sqrt{n}}{n_s\sqrt{\log n} \cdot \sqrt{k}} \cdot W)$ when $n_s \to \infty$. Since k is constant in this case, we know that the per-flow multicast capacity is upper-bounded by the per-flow unicast capacity with n_s unicast sessions. Thus, the per-flow multicast capacity is almost surely at most $O(\frac{n}{n_s\sqrt{\log n}} \cdot W)$, which is same as $O(\frac{n}{n_s\sqrt{\log n} \cdot \sqrt{k}} \cdot W)$ since k is constant.

C. When $k = \Omega(\frac{a^2}{r^2})$

In the previous subsection, we showed an upper bound of the multicast capacity when $k < \theta_1 \cdot a^2/r^2$. In this subsection we will present an upper bound on multicast capacity when $k \geq \theta_1 \cdot a^2/r^2$. We will essentially show that in this case, multicast is asymptotically equivalent to broadcast. Broadcast capacity of *single-source* of an arbitrary network has been studied in [9], [18]. In this paper, we will prove that the achievable integrated multicast capacity is only $\Theta(W)$ if an arbitrary k subset of the n nodes will serve as receivers for each possible source node v_i.

We partition the square of side-length a into squarelets, each with side length r. The square will be partitioned into $M = \lceil a^2/r^2 \rceil$ squarelets, say $B_1, B_2, \cdots, B_M$. Recall that we will randomly select $k \geq \theta_1 \cdot a^2/r^2$ receivers in the square region. See Figure 1 (b) for illustration.

Lemma 17: With probability at least $1 - \frac{1}{\rho^2 e^{\theta_1} M}$, at least $\rho \cdot M$ squarelets will have at least one receiver when $k \geq \theta_1 \cdot a^2/r^2$ for a constant θ_1.

Proof: Let X be the number of squarelets that do not have any receivers inside, and A be a fixed fraction of squarelets, say $A = \rho \cdot M$ for a constant $0 < \rho < 1$. Define variable X_i

$$X_i = \begin{cases} 1 & \text{if squarelet } B_i \text{ is empty of receivers,} \\ 0 & \text{if squarelet } B_i \text{ is not empty of receivers.} \end{cases} \quad (10)$$

Notice $X = \sum_{i=1}^{n} X_i$, and $\text{Var}(X) = \text{Var}(\sum_{i=1}^{M} X_i) = \sum_{i=1}^{M}\sum_{j=1}^{M} \text{Cov}(X_i, X_j)$, where $\text{Cov}(X_i, X_j) = E(X_i \cdot X_j) - E(X_i)E(X_j)$ is the covariance of variable X_i and X_j. We then compute such $\text{Cov}(X_i, X_j)$ for all possible pairs of i and j: $\text{Cov}(X_i, X_i) = E(X_i) - E(X_i)^2$ and $E(X_i) = (1 - \frac{1}{M})^k$; and $E(X_i \cdot X_j) = (1 - \frac{2}{M})^k$ if $i \neq j$. Consequently, we have:

$$\begin{aligned}\text{Var}(X) &= M(M-1)[(1-\frac{2}{M})^k - (1-\frac{1}{M})^{2k}] \\ &\quad +M[(1-\frac{1}{M})^k - (1-\frac{1}{M})^{2k}] \\ &= M(M-1)[(1-\frac{2}{M})^k - (1-\frac{2}{M}+\frac{1}{M^2})^k] \\ &\quad +M[(1-\frac{1}{M})^k - (1-\frac{1}{M})^{2k}]\end{aligned}$$

Since $[(1-\frac{2}{M})^k - (1-\frac{2}{M}+\frac{1}{M^2})^k] \leq 0$, we have

$$\text{Var}(X) \leq M[(1-\frac{1}{M})^k - (1-\frac{1}{M})^{2k}].$$

From Lemma 1, we have

$$\mathbf{Pr}\,(X - E(X) \geq \rho \cdot M) \leq \frac{M[(1-\frac{1}{M})^k - (1-\frac{1}{M})^{2k}]}{\rho^2 M^2}$$

Recall that $k \geq \theta_1 \cdot M$. The expected value $E(X)$ of X, denoted by μ, is $M \cdot (1-\frac{1}{M})^k \leq M \cdot e^{-\theta_1}$. Thus,

$$\mathbf{Pr}\left(X \geq (e^{-\theta_1} + \rho) \cdot M\right) \leq \frac{(\frac{1}{e})^{\theta_1} - (\frac{1}{e})^{2\theta_1}}{\rho^2} \cdot \frac{1}{M}$$

When $M \to \infty$, the probability goes to zero. We can also show that, with high probability, there is at most a constant fraction of squarelets that will be empty of receivers. This finishes the proof. ■

We then prove that the union of the transmission disks of these k nodes ($k-1$ receivers and 1 source node) in a multicast will cover at least a constant fraction, say $0 < \rho_2 \leq 1$, of the deployment region.

Lemma 18: The union of the transmission disks of these k nodes ($k-1$ receivers and 1 source node) in a multicast will cover at least a constant fraction, say $0 < \rho_2 \leq 1$, of the deployment region.

Proof: Based on lemma 17, we know that among M squarelets partitioned from the deployment region, there are at least $\rho \cdot M$ squarelets, each of which contains at least one receiver (or source) node inside. In each such a squarelet B_j, there is at least one receiver and thus at least one transmitting node in the multicast tree that covers this receiver. The transmitting node must lie inside this squarelet or 8 adjacent squarelets. On the other hand, each transmitting disk can cover receivers from at most 9 squarelets. Consequently, we must have at least $\rho \cdot M/9$ transmitting disks to cover receivers from $\rho \cdot M$ squarelets. Recall that the squarelet side-length is r, which implies that each point in the deployment region is covered by *at most* 9 such representative transmission disks. Consequently, the total area covered by these representative transmission disks is at least $\rho \cdot M \cdot \pi r^2/81$. Recall that the deployment region has area a^2 and $M = \lceil a^2/r^2 \rceil$. Thus, the area of all transmission disks of all these k nodes is at least $\rho_2 = \frac{\rho \cdot \pi}{81}$ fraction of the total area of the deployment region. This finishes the proof. ■

Based on the above lemmas, the following theorem is straightforward.

Theorem 19: When $k \geq \theta \cdot a^2/r^2$ for a constant θ, the total multicast capacity $\Lambda_k(n)$ of all nodes is bounded from above by

$$\Lambda_k(n) \leq \frac{W \cdot a^2}{\rho_2 a^2} = \frac{W}{\rho_2} = O(W)$$

where ρ_2 is a constant depending only on θ.

Proof: The total multicast capacity of all nodes is bounded from above by

$$\frac{W \cdot a^2}{\rho_2 a^2} = \frac{W}{\rho_2}.$$

Notice that here $W \cdot a^2$ is the total bits×meter^2/sec that can be occupied by all nodes' transmissions: the transmission of one node will cover an area at most πr^2 and the transmission disks of active transmitting nodes at any time instance should be disjoint, which implies that, the total bits×meter^2/sec achievable is at most $W \cdot a^2$. On the other hand, for each bit in a multicast session, it must cover at least $\rho_2 a^2$ meter^2 in the deployment region. This finishes the proof. ■

Notice that for broadcast, it has been proved in [9], [18] that the broadcast capacity is only $\Theta(W)$. Here we essentially prove that for multicast, when the number of receivers is large enough (at least $\Omega(\frac{a^2}{r^2})$), the asymptotic multicast capacity is also only $O(W)$.

V. Lower Bounds on Multicast Capacity with Random Networks

In the previous section, we have derived upper bounds on the multicast capacity $\Lambda_k(n)$. In this section, we will derive asymptotically matching lower bounds on the multicast capacity $\Lambda_k(n)$. Specifically, we will provide a multicast scheme and prove that the multicast capacity achieved by our scheme matches the asymptotic upper bounds.

A. Partition Square Using Squarelets

Our multicast scheme is based on a good approximation of a minimum connected dominating set (MCDS) of a random network. First, partition the region into squarelets, each of side-length $r/\sqrt{5}$. Thus, any two nodes from 2 adjacent squarelets (sharing a common side) will be able to communicate with each other directly. Randomly select one node from each squarelet. Clearly the set of selected nodes is a dominating set. If every squarelet has a node inside it, obviously, the set of selected nodes will form a connected dominating set (CDS).

Notice that, it is possible that, for some squarelet, there is no node inside, and thus, we cannot find a multicast tree $MT(U_1')$ later by Algorithm 1. We show that this almost surely cannot happen.

Lemma 20: There is a sequence of $\delta(n) \to 0$ such that

$$\mathbf{Pr}\,(\text{Every squarelet contains a node}) \geq 1 - \delta(n)$$

Proof: Let $\mathcal{C}$ be the class of axis-aligned squares of side-length $\frac{r}{\sqrt{5}}$. Notice that the probability that a node fall in such a square is $\frac{r^2}{5} \cdot \frac{1}{a^2} = \frac{r^2}{5a^2}$. Recall that, to have a connected network, we almost surely have $r/a \geq \sqrt{\frac{\log n}{n\pi}}$. It is easy to show that the VC-dimension of $\mathcal{C}$ is at most 4 (it is at least 3, [4]). Hence, for all squarelets S,

$$\mathbf{Pr}\left(\sup_{S\in\mathcal{C}}\left|\frac{\#\text{ of nodes in } S}{n} - \frac{r^2}{5a^2}\right| \leq \epsilon(n)\right) > 1-\delta(n)$$

whenever

$$n \geq \max\left\{\frac{32}{\epsilon(n)}\cdot\log\frac{13}{\epsilon(n)}, \frac{4}{\epsilon(n)}\log\frac{2}{\delta(n)}\right\}. \quad (11)$$

This condition 11 is satisfied when

$$\epsilon(n) = \frac{32\log n}{n}, \delta(n) = \frac{2}{n}.$$

Thus,

$$\mathbf{Pr}\left(\sup_{S\in\mathcal{C}}\{\#\text{ of nodes in } S \geq \frac{nr^2}{5a^2} - n\cdot\epsilon(n)\}\right) > 1-\delta(n)$$

Thus, if we choose r and a such that

$$r \geq 14\cdot a\cdot\sqrt{\frac{\log n}{n}} \quad (12)$$

then $\frac{nr^2}{5a^2} - n\cdot\epsilon(n) \geq 7\log n$. Consequently, we have

$$\mathbf{Pr}\left(\forall\text{ squarelet } S, \#\text{ of nodes in } S \geq 7\log n\right) > 1-\frac{2}{n}$$

The theorem then follows. ■

Notice that, generally, when $r = a\sqrt{\frac{\log n}{c\cdot n}}$, to make sure that every squarelet with side-length r/β will have at least one node inside, it is sufficient to require that $c < \frac{1}{32\beta^2}$.

For each node on the CDS, we show that every node can be scheduled to transmit once every Δ time-slots, where constant Δ depending only on R and r. For each node v, consider a node u whose transmission will interfere with the transmission of node v. Clearly node u will be completely inside the disk centered at v with radius $R+r$. Thus, the squarelet containing u must be inside the disk centered at v with radius $R+r+\frac{\sqrt{2}}{\sqrt{5}}r < R+2r$. Let Δ be the maximum number of nodes in CDS whose transmission will interfere with the transmission of a node v in CDS. Using the area argument, we can show that

$$\Delta \leq \frac{\pi\cdot(R+2r)^2}{r^2/5} = 5\pi(2+\frac{R}{r})^2.$$

This property ensures that we can schedule the transmissions of all nodes in CDS by a TDMA manner such that all nodes will be able to transmit at least once in every Δ time slots. Notice that here Δ is a constant.

[4]A detailed analysis can show that its VC-dimension is exactly 3. Consider any four points x_1, x_2, x_3 and x_4. If the convex hull of them has 3 corners, then one node, say x_4 is inside. Obviously, any square containing x_1, x_2 and x_3 will have x_4 inside. So $\{x_1, x_2, x_3\}$ cannot be realized. When the convex hull has all 4 nodes, let $x_1x_2x_3x_4$ be the convex hull in clockwise order. Since $\{x_1, x_2, x_3, x_4\}$ can be realized, they must be inside a square with the side-length $r/\sqrt{5}$. Then it is easy to show that either $\{x_1, x_3\}$ or $\{x_2, x_4\}$ cannot be realized. This finishes the proof.

B. When $k \leq \theta_1 a^2/r^2$

When the number of receivers, plus the source node, k is at most $\theta_1\frac{a^2}{r^2}$, we will construct a multicast tree from CDS. Consider an instance of a random network $G = (V, E)$ and also an instance of multicast with v_1 as the source node and $U_1 = \{v_2, v_3, \cdots v_k\}$ as the receiver nodes. Let $U_1' = \{v_1, v_2, v_3, \cdots v_k\}$. We will construct a multicast structure as following Algorithm 1.

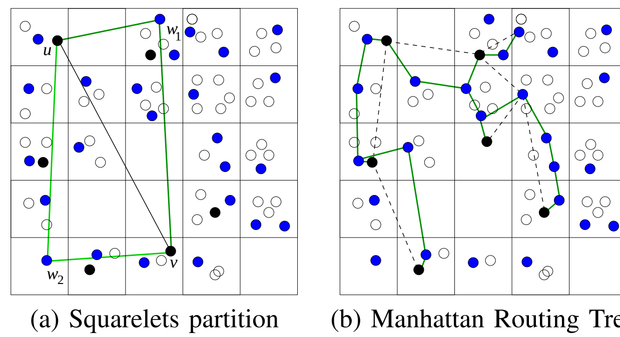


Fig. 2. Partition of square with side-length a into squarelets with side-length $r/\sqrt{5}$. For an edge $uv \in EMST(U_1')$, find a node w (either node w_1 or node w_2 which has same row as u and same column as v) to connect them. Right figure illustrates a multicast tree constructed using Manhattan approach, where dotted lines denote an original spanning tree of nodes in a multicast session.

To show that the above routing Algorithm 1 achieves the asymptotically optimum multicast capacity, we need show that the total number of copies that a multicast bit is "received" by all nodes is at most $O(\frac{r\cdot\sqrt{k}\cdot n}{a})$, which will be derived based on the upper bound on the area covered by all transmission disks in the multicast tree $MT(P_i)$. The following theorem will bound total Euclidean length of edges in the multicast tree $MT(P_i)$

Theorem 21: The total Euclidean length of the multicast tree $MT(P_i)$ is at most $c_3\sqrt{k}a$ for a constant c_3 depending only on θ_1.

Proof: For k nodes P_i in a multicast session with source node v_i, we first constructed the Euclidean spanning tree $EST(P_i)$. Same as the proof for Lemma 12, the total Euclidean length of all edges in $EST(P_i)$ is at most $2\sqrt{2}\sqrt{k}a$. For each edge $uv \in EST(P_i)$, we will select a sequence of nodes (one in each of the squarelets crossed by segment uw or wv) to connect them. Assume that segment uw crosses x squarelets (including the one containing u and w), and segment wv crosses y squarelets. Let s be the size of the squarelet, which is $r/\sqrt{5}$ here. Then the path $\mathbf{P}(u, v)$ has Euclidean length at most $(x+y)s\cdot\sqrt{2}$ because the path has length at most $\sqrt{2}s$ in each crossed squarelet and it will cross $x+y$ squarelets. Additionally, we have

$$\begin{cases}(x\cdot s)^2 + (y\cdot s)^2 \geq \|uv\|^2\\((x-1)\cdot s)^2 + ((y-1)\cdot s)^2 \leq \|uv\|^2\end{cases} \quad (13)$$

Thus, $(x-1)+(y-1) \leq \sqrt{2}\sqrt{(x-1)^2+(y-1)^2} = \sqrt{2}\frac{\|uv\|}{s}$. Then, $\mathbf{P}(u, v)$ has Euclidean length at most

$$(x+y)s\cdot\sqrt{2} \leq 2\|uv\| + 2\sqrt{2}s.$$

Algorithm 1 Multicast Capacity Achieving Manhattan Routing Based on a Squarelet for Nodes U_i

1: We partition the deployment square into squarelets, each with side length $r/\sqrt{5}$ (as in [16], see Figure 1 (b) for illustration). Thus, we have $\lceil \frac{a}{r/\sqrt{5}} \rceil$ squarelets. Each squarelet is denoted by (i, j) when it is the ith column and jth row.

2: Let $P_i = \{p_{i,1}, p_{i,2}, \cdots, p_{i,k}\}$ be the set of randomly and independently selected points used to find the terminals $U_i = \{v_{i,1}, v_{i,2}, \cdots, v_{i,k}\}$. Recall that $v_{i,j}$ is the closest node to point $p_{i,j}$.

3: We build an Euclidean spanning tree, denoted as $EST(P)$, connecting points in P_i, using following method (also described in Lemma 12):
(1) Originally, k nodes P_i form k components;
(2) repeat steps (3) and (4) for $g = 1, 2, \cdots, k-1$,
(3) for the gth step, partition the deployment square into at most $k-g$ square-shaped-cells, each with side length $\lceil \frac{a}{\lfloor \sqrt{k-g} \rfloor} \rceil$;
(4) find a cell that contains two points of P_i that are from 2 different connected components and then connect them using Manhattan routing; merge these two connected components.

4: For each link uv in the tree $EST(P_i)$, assume that u and v are inside squarelet (i_u, j_u) and squarelet (i_v, j_v) respectively. Find a point w in squarelet (i_v, j_u) (or squarelet (i_u, j_v)), *i.e.*, uwv is a Manhattan path connecting u and v. See Figure 2 (a) for illustration. The resulted structure by uniting all such paths for all links in $EST(P_i)$ will serve the routing guideline for multicast.

5: For each edge uw in $EST(P_i)$, find a node in each of the squarelets that are crossed by line uw. We connect these nodes in sequence to form a path, denoted as $\mathbf{P}(u, v)$, connecting points u and v. Notice that here such structure may not be a tree. If this is the case, we could remove the cycles that do not contain nodes from P_i. Denote the resulted tree as $MT(P_i)$.

6: For each receiver $v_{i,j}$, if it is not inside the squarelet $\mathbf{s}$ containing point $p_{i,j}$, let $v'_{i,j}$ be the node selected inside the squarelet $\mathbf{s}$. Notice that, such $v'_{i,j}$ exists for every squarelets, with probability at least $1 - 2/n$. Node $v'_{i,j}$ then relay the data to node $v_{i,j}$ (the relay takes at most 2 hops). The final tree (including these additional relays) are called multicast tree $MTR(U_i)$.

The total Euclidean length of all edges in $MT(P_i)$ is at most

$$\begin{aligned} & \sum_{uv \in EST(P_i)} 2\|uv\| + 2\sqrt{2}s \\ = \;& 2\|EST(P_i)\| + 2\sqrt{2}\frac{r}{\sqrt{5}} \cdot k \\ \leq \;& 2 \cdot (2\sqrt{2}\sqrt{k}a) + \frac{2\sqrt{10}}{5}\sqrt{k} \cdot a \cdot \frac{r}{a} \cdot \sqrt{k} \\ \leq \;& (4\sqrt{2} + \frac{2\sqrt{10}}{5} \cdot \sqrt{\theta_1})\sqrt{k} \cdot a. \end{aligned}$$

The last inequality comes from $k \leq \theta_1 a^2/r^2$. This finishes proof by setting $c_3 = 4\sqrt{2} + \frac{2\sqrt{10}}{5} \cdot \sqrt{\theta_1}$. ∎

Here we denote the region covered by all transmission disks of all internal nodes in $MT(P_i)$ as $D(T)$ and the number of nodes lying in $D(T)$ as C. We then show that with high probability, the multicast capacity achieved using above routing approach is within a constant factor of the asymptotic optimum. We essentially show that, with high probability, the number C of nodes that will receive a *copy* of the multicast data is within $2E(C)$.

Lemma 22: Given a multicast tree constructed by Algorithm 1, the number of nodes that will get a copy of a multicast data is, with high probability, at most $c_4 \frac{n \cdot r \cdot \sqrt{k}}{a}$, when $k \leq \theta_1 \frac{a^2}{r^2}$. Here c_4 is a constant.

Proof: Consider a set of receivers U_1 for source node v_1. Let tree T be the multicast tree $MT(P_i)$ constructed above. Let $X_i \in \{0, 1\}$ be an indicator variable whether the ith node v_i will fall inside the region $D(T)$ for a multicast tree T. Clearly $p = \mathbf{Pr}(X_i = 1) = \frac{|D(T)|}{a^2}$. Notice that the area of $D(T)$ is at most $2r \cdot \|T\| + k\pi r^2/2$, and edge length $\|T\| \leq c_3 \cdot \sqrt{k} \cdot a$. Obviously, $X = \sum_{i=1}^{n} X_i$ is the number of nodes falling inside the region $D(T)$, and X is binomial distribution. Using Lemma 4, we have

$$\begin{aligned} & \mathbf{Pr}\left(C > |D(T)| \cdot \frac{2n}{a^2}\right) \leq \frac{|D(T)| \cdot \frac{2n}{a^2} \cdot (1 - \frac{|D(T)|}{a^2})}{(|D(T)| \cdot \frac{2n}{a^2} - \frac{n \cdot |D(T)|}{a^2})^2} \\ = \;& \frac{2[1 - \frac{|D(T)|}{a^2}]}{|D(T)| \cdot \frac{n}{a^2}} \leq \frac{2a^2}{n \cdot |D(T)|} \leq \frac{2c_0 \cdot a^2}{n \cdot \tau\sqrt{k} \cdot a \cdot r} \\ = \;& \frac{a}{r} \frac{1}{n\sqrt{k}} \frac{2c_0}{\tau} \leq \frac{1}{\sqrt{n \cdot k \cdot \log n}} \frac{2c_0\sqrt{c}}{\tau} \end{aligned}$$

The last inequality comes from the assumption that $a/r \leq \sqrt{\frac{c \cdot n}{\log n}}$. The second to last inequality comes from Lemma 13 that $|D(T)| \geq \frac{\tau\sqrt{k} \cdot a \cdot r}{c_0}$, *w.h.p.*. Consequently,

$$\mathbf{Pr}\left(C \leq |D(T)| \cdot \frac{2n}{a^2}\right) \geq 1 - \frac{1}{\sqrt{n \cdot k \cdot \log n}} \frac{2c_0\sqrt{c}}{\tau}.$$

Thus the number of nodes that can get a copy of the data for multicast within nodes P_i, with high probability, is at most

$$|D(T)| \cdot \frac{2n}{a^2} \leq c_3\sqrt{k} \cdot \frac{4n \cdot r}{a} + \pi n k \frac{r^2}{a^2} \leq (4c_3 + \pi\sqrt{\theta_1}) \cdot n\sqrt{k} \cdot \frac{r}{a},$$

The last inequality comes from $k \leq \theta_1 \frac{a^2}{r^2}$. The lemma follows by setting $c_4 = 4c_3 + \pi\sqrt{\theta_1}$. ∎

Recall that, by performing multicast based on CDS structure, we can guarantee that each node will be able to transmit once every Δ time-slots. This implies that the total bits/sec achieved by all nodes is at least $n \cdot W/\Delta$. Consequently, the multicast capacity is at least

$$\frac{n \cdot W/\Delta}{(4c_3 + \pi\sqrt{\theta_1}) \cdot n\sqrt{k} \cdot \frac{r}{a}} = \frac{1}{(4c_3 + \pi\sqrt{\theta_1}) \cdot \Delta} \cdot \frac{a \cdot W}{r\sqrt{k}}$$

By setting, $c_5 = \frac{1}{(4c_3 + \pi\sqrt{\theta_1}) \cdot \Delta}$, we have the following theorem.

Theorem 23: The total multicast capacity $\Lambda_k(n)$ achievable by all multicast flows is at least $c_5 \frac{a \cdot W}{r\sqrt{k}}$, when $k \leq \theta_1 \frac{a^2}{r^2}$ and

$a/r \leq \sqrt{\frac{cn}{\log n}}$ for some constant $c \in (0, 1/160]$. Here c_5 is a constant.

Observe that the correctness of Theorem 23 relies on the fact that $\frac{a}{r} \leq \sqrt{\frac{cn}{\log n}}$ and $k \leq \theta_1 a^2/r^2$. Here constant $0 < c < 1/160$, from condition (12). Consequently, by letting $\frac{a}{r} = \sqrt{\frac{cn}{\log n}}$ for $0 < c < 1/160$, and $c_2' = c_5\sqrt{c}$, based on Theorem 23, we have

Corollary 24: The multicast capacity for a random network of n nodes, when $k < \theta_1 \cdot a^2/r^2$, is at least

$$\Lambda_k(n) \geq c_2' \cdot \frac{\sqrt{n}}{\sqrt{\log n} \cdot \sqrt{k}} \cdot W = \Omega(\frac{\sqrt{n}}{\sqrt{\log n} \cdot \sqrt{k}} \cdot W).$$

Consequently, the multicast capacity per flow (with n sources) is at least

$$\lambda_k(n) = \frac{\Lambda_k(n)}{n} = \Omega(\frac{1}{\sqrt{n \log n} \cdot \sqrt{k}} \cdot W).$$

Observe that the correctness of Theorem 23 requires the following two additional properties of our routing scheme

1) The traffic "load" of *every* routing squarelet is no more than a constant factor of W bits/sec (at most $\frac{W}{\Delta}$ in this paper), due to the requirement of TDMA node scheduling.
2) At least a constant fraction of nodes V that will send the multicast data or is within the transmission range of some transmitting nodes. This is required for using nW/Δ as an approximation of the total bits "received" by all nodes per unit time. This condition is clearly satisfied when every node could serve as the multicast source. We will prove that it is still true when there are n_s multicast sessions and n_s satisfies some condition.

We will first prove that, for any squarelet $\mathbf{s}$, with high probability, the *traffic load* (total data rates) assigned to nodes in $\mathbf{s}$ is at most W/Δ. Given a squarelet, we define its *flow-load* as the *total number* of multicast sessions that will be routed through nodes inside this squarelet. We show that under our routing algorithm, for any squarelet, with high probability, its flow-load is no more than $\Theta(\sqrt{kn \log n})$. To prove our claim, we first study a simple unicast case. Consider a grid of $L \times L$ squarelets. Consider a specific squarelet $\mathbf{s}$ that is of ith row and jth column in the squarelet-grid. See Figure 3 for an illustration. Randomly pick two nodes u and v from

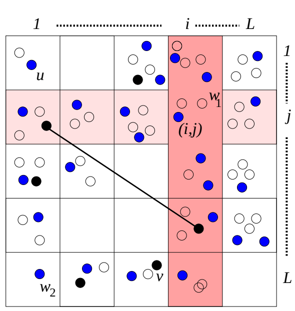


Fig. 3. The expected flow load on a squarelet.

the grid and connect them via Manhattan routing. Let $X_{\mathbf{s}}$ be a random variable denoting whether the Manhattan routing will use nodes from the squarelet $\mathbf{s}$. Let $p_{\mathbf{s}}(L)$ denote the probability that the Manhattan routing will use nodes from the squarelet $\mathbf{s}$, *i.e.*, $p_{\mathbf{s}}(L) = \mathbf{Pr}(X_{\mathbf{s}} = 1)$. Then

$$p_{\mathbf{s}}(L) = \frac{i-1}{L^2} \cdot \frac{L-i+1}{L} + \frac{j-1}{L^2} \cdot \frac{L-j+1}{L}. \quad (14)$$

Here $\frac{i-1}{L^2} \cdot \frac{L-i+1}{L}$ (resp. $\frac{j-1}{L^2} \cdot \frac{L-j+1}{L}$) is the probability that squarelet $\mathbf{s}$ is used when u (resp. v)is on the same row (resp. column) as $\mathbf{s}$. It is easy to show that

$$\frac{2}{L^2} \leq p_{\mathbf{s}}(L) \leq \frac{2}{L}.$$

Let us now study the number of times that a specific squarelet $\mathbf{s}$ is used by our routing structure for multicast.

Lemma 25: Given a squarelet $\mathbf{s}$, the probability that a random multicast flow will be routed via the squarelet $\mathbf{s}$ is at most $c_6\sqrt{k} \cdot \frac{r}{a}$.

Proof: Recall that we will construct the Euclidean spanning tree as the method described in Algorithm 1 and then find multicast routing structure as Algorithm 1. For a given multicast session, this squarelet $\mathbf{s}$ may be used in any one of the k steps to build the spanning tree. For step g (with $1 \leq g \leq k-1$), recall that we will partition the square with side-length a into $\lfloor\sqrt{k-g}\rfloor^2 \leq k-g$ cells, each with side-length $\frac{a}{\lfloor\sqrt{k-g}\rfloor}$. From pigeonhole principle, there exists a cell that contains two nodes, say u and v, from two different connected components. We will connect them and merge these two connected components. Here we will connect u and v using Manhattan routing as illustrated in Figure 2. Let $X_{\mathbf{s},g}$ be the indicator whether the specific squarelet $\mathbf{s}$ is used in this gth step. Clearly, the probability that $\mathbf{Pr}(X_{\mathbf{s},g} = 1)$ is

$$\mathbf{Pr}(X_{\mathbf{s},g} = 1) = \frac{1}{\lfloor\sqrt{k-g}\rfloor^2} \cdot p_{\mathbf{s}}(\lceil\frac{\frac{a}{\lfloor\sqrt{k-g}\rfloor}}{r/\sqrt{5}}\rceil), \quad (15)$$

where $\frac{1}{\lfloor\sqrt{k-g}\rfloor^2}$ is the probability that the cell containing squarelet $\mathbf{s}$ is used, and $p_{\mathbf{s}}(\lceil\frac{\frac{a}{\lfloor\sqrt{k-g}\rfloor}}{r/\sqrt{5}}\rceil)$ is the probability that squarelet $\mathbf{s}$ is used when that cell containing $\mathbf{s}$ is used. Here $\lceil\frac{\frac{a}{\lfloor\sqrt{k-g}\rfloor}}{r/\sqrt{5}}\rceil$ is the number of squarelets per row in a cell, *i.e.*, the value of L in formula 14. Consequently,

$$\begin{aligned}
p &= \mathbf{Pr}(X_{\mathbf{s}} = 1) \leq \sum_{g=1}^{k-1} \mathbf{Pr}(X_{\mathbf{s},g} = 1) \\
&= \sum_{g=1}^{k-1} \frac{1}{\lfloor\sqrt{k-g}\rfloor^2} \cdot p_{\mathbf{s}}(\lceil\frac{\frac{a}{\lfloor\sqrt{k-g}\rfloor}}{r/\sqrt{5}}\rceil) \\
&\leq \sum_{g=1}^{k-1} \frac{1}{k-g} \cdot p_{\mathbf{s}}(\lceil\frac{\sqrt{5}a}{r\lfloor\sqrt{k-g}\rfloor}\rceil) \\
&\leq \sum_{g=1}^{k-1} \frac{1}{k-g} \cdot 2 \cdot \frac{r\sqrt{k-g}}{\sqrt{5}a} \\
&= \sum_{g=1}^{k-1} \frac{1}{\sqrt{g}} \cdot \frac{2r}{\sqrt{5}a} \\
&\leq \frac{4\sqrt{10}}{5}\sqrt{k} \cdot \frac{r}{a}
\end{aligned}$$

The theorem follows from by setting $c_6 = \frac{4\sqrt{10}}{5}$. ∎

Similarly, using $\frac{2}{L^2} \leq p_{\mathbf{s}}(L)$, we can show that

Lemma 26: Given a squarelet **s**, the probability that a random multicast flow will be routed via the squarelet **s** is at least $k \cdot \frac{r^2}{5a^2}$.

Thus, given any squarelet **s**, the expected number of flows that will be routed through the squarelet **s** is at most $c_6 \cdot n_s\sqrt{k}\frac{r}{a}$, given n_s multicast sessions. Notice that, to achieve larger multicast capacity, we will set $\frac{a}{r} = \sqrt{\frac{cn}{\log n}}$ for some constant $0 < c \leq 1/160$ (see proof of Lemma 20). Thus, $\mathbf{Pr}\left(X_{\mathbf{s}} = 1\right) \leq c_6\sqrt{\frac{k \cdot \log n}{cn}}$. Then we have the following lemma

Lemma 27: Given n_s multicast sessions, the expected number of multicast routing flows that use a specific squarelet **s** is at most $\frac{c_6}{\sqrt{c}} \cdot n_s \cdot \sqrt{\frac{k \cdot \log n}{n}}$. When $n_s = n$, it is at most $\frac{c_6}{\sqrt{c}} \cdot \sqrt{k \cdot n \cdot \log n}$.

Recall that, the multicast rooted at v_i will first randomly and independently select $k-1$ points P_i'. To use the VC Theorem, we will construct a multicast tree using the union of node v_i and P_i' as P_i, which is the input of Algorithm 1. The data will then be relayed to every node $v_{i,j}$ if it did not receive the data before. Thus, the *points* used to construct the multicast trees $MT(P_i)$ for different source nodes are independently and randomly chosen in the deployment region. Notice that, given k terminals U, the multicast tree $MT(U)$ constructed by Algorithm 1 can be uniquely defined by its terminals U, thus has dimension $2k$. In other words, every point in $2k$-dimensional cube (with side-length a), corresponds to a multicast tree. Given any 2-dimensional axis-aligned square h (not necessarily the squarelet produced by partitioning the deployment region), let set $T(h)$ be the set of multicast trees (equivalently, the set of points in R^{2k} defining these trees) that will intersect the square h (*i.e.*, one of its edges will have point inside h). Let

$$\mathcal{F} = \{T(h) \mid h \text{ is an axis-aligned square with size } \frac{r}{\sqrt{5}}\}.$$

We will show that the VC-dimension of $\mathcal{F}$ is at most $d = \Theta(\log k)$.

To prove this, we first study the VC-dimension of the following system. Let X be the universal set of 2-dimensional segments. For an axis-aligned square h with a fixed side-length, let $X(h)$ be the set of all segments from X that intersect (or is contained inside) the square h. Let

$$\mathcal{S} = \{X(h) \mid h \text{ is an axis-aligned square with size } \frac{r}{\sqrt{5}}\}.$$

Given *any* set of m line segments $\mathcal{L} = \{L_1, L_2, \cdots, L_m\}$, we show that the cardinality of

$$\Pi_{\mathcal{S}}(\mathcal{L}) = \{\mathcal{L}(h) \mid h \text{ is a 2D square with side-length } \frac{r}{\sqrt{5}}\}$$

is polynomial of m.

Lemma 28: The cardinality of $\Pi_{\mathcal{S}}(\mathcal{L})$ is at most $2m^2$, where m is the cardinality of $\mathcal{L}$.

Proof: This is essentially to study the number of different sets of segments that can be intersected by *all* 2-dimensional solid square h. Imagine that we move a square J (with fixed side-length) all over the 2-dimensional space and at any moment we can only see the region not covered by the square J. A *view* is defined as the set of (partial or full) segments that can be seen through the square J. Then the view will change only if one the following 4 events happens:

1) A segment starts entering the view and the first point seen is its end-point.
2) A segment starts entering the view and the first point seen is an interior point of the segment.
3) A segment starts leaving the view and the last point seen is its end-point.
4) A segment starts leaving the view and the last point seen is an interior point of the segment.

For scenarios 1) and 3), the square J must have one of its side-edge touching an end-point of some segment. For scenarios 2) and 4), it must be the case that a corner of the square J touches the segment on that interior point. See Figure 4 for illustration. When one of the above 4 events happens to the square J, we

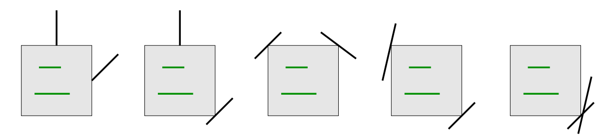


Fig. 4. Scenarios when a square is stable.

call that J has a *support*. Obviously, when a square has only one *support*, we can still move the square without changing the view, while keep this support. We can keep moving until the square has another *support*. A square (axis-aligned and with a fixed side-length) is called *stable* if it has at least 2 supports (at least one at x-axis and one at y-axis). Notice that there is a degenerate case here: we cannot find another support when we move a square to find another support. In this degenerated case, the square is called stable if one of its corners touches an end-point of a segment. Consequently, the cardinality of $\Pi_{\mathcal{S}}(\mathcal{L})$ is at most the number of stable squares produced by these segments. Notice that, given any pair of segments, it can produce at most 4 stable squares. See Figure 5 for illustration. Thus, the cardinality of $\Pi_{\mathcal{S}}(\mathcal{L})$ is at most $\frac{m^2}{2} \cdot 4 = 2m^2$. This

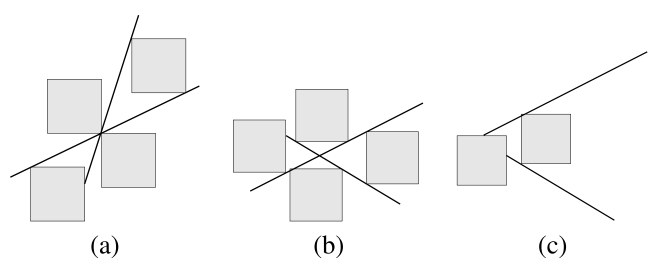


Fig. 5. Scenarios when two segments can produce a stable square.

finishes the proof. ■

Notice that, for m segments $\mathcal{L}$, the cardinality of $\Pi_{\mathcal{S}}(\mathcal{L})$ is at most $\frac{m^2}{2} \cdot 4 = 2m^2$. It implies that when a set $\mathcal{L}$ with cardinality m is shatterable by $\mathcal{S}$, $2m^2 \geq 2^m$. Thus, $m < 7$. Consequently, the VC-dimension of $\mathcal{S}$ is at most 6.

Lemma 29: The VC-dimension of $\mathcal{F}$ is at most $d = O(\log k)$.

Proof: Consider any m trees $\mathcal{T} = \{T_1, T_2, \cdots, T_{m-1}, T_m\}$ that is shatterable by $\mathcal{F}$. If we consider only the segments in these trees, there are $m(k-1)$ segments. From Lemma 28, we know that the cardinality of $\Pi_{\mathcal{F}}(\mathcal{T})$ is at most $2m^2 \cdot (k-1)^2$. Thus, if $\mathcal{T}$ is shatterable by $\mathcal{F}$, we have

$$2m^2 \cdot (k-1)^2 \geq 2^m$$

Thus, $m < 3 \log k$, when $k \geq 2^{12}$. When $k < 2^{12}$, m is also at most a constant. This finishes the proof. ■

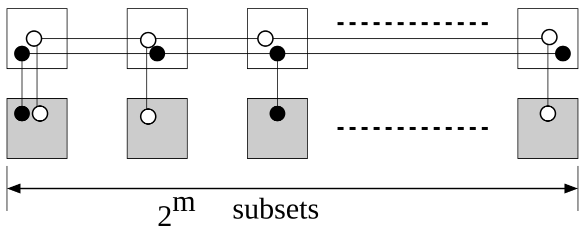


Fig. 6. An example that VC-dimension of trees is $\Theta(\log k)$. Here shaded squares will produce 2^m distinctive subsets. Here white nodes belong to one tree, and black nodes belong to another tree.

Lemma 30: The VC-dimension of $\mathcal{F}$ is at least $d = \Omega(\log k)$.

Proof: We show by example that the VC-dimension of $\mathcal{F}$ is at least $\log \frac{2k}{3}$. See Figure 6 for illustration. In the example, we will present m trees such that we can view exactly any subset of these m trees using a shaded square with a fixed side-length ℓ. There are 2^m such subsets. We will have 2^m disjoint shaded squares, each corresponding to a subset. Every tree will put one terminal in a shaded square if it belongs to the corresponding subset. Thus, we need exactly 2^{m-1} terminals from each tree that will be put in these shaded squares. To make sure that trees generated will *not* cross a shaded square which it does not belong to, we will an add additional node on top of each shaded square: in the spanning tree, a node in the shaded square will *only* connect with nodes in the unshaded squares. Then it is easy to see that the total number of nodes a tree needed is $2^m + 2^{m-1}$ (each for 2^m squares on top and each for 2^{m-1} squares on bottom). Thus, we set m as $2^{m+1} + 2^m > k \geq 2^m + 2^{m-1}$. This implies that $\log \frac{2k}{3} - 1 < m \leq \log \frac{2k}{3}$. The theorem then follows. ■

Consequently, the VC-dimension of $\mathcal{F}$ is $d = \Theta(\log k)$.

Theorem 31: Assume that there are N random multicast sessions. There is a sequence of $\delta(n) \to 0$ such that

$$\begin{aligned}&\mathbf{Pr}\left(\forall \text{ squarelet } \mathbf{s}, \# \text{ of flows using } \mathbf{s} \leq \frac{3\sqrt{c_6}N}{2}\sqrt{k}\frac{r}{a}\right)\\ \geq\ & 1 - \delta(n)\end{aligned}$$

Proof: The terminals to constructed multicast trees are *i.i.d.* variables. Then the multicast trees are *i.i.d.* variables. Thus, we can use the VC-Theorem. Recall that, given a square h, the probability that a multicast tree will cross h is at most $c_6\sqrt{k} \cdot \frac{r}{a}$ (see Lemma 25). Hence, for all squarelets S,

$$\mathbf{Pr}\left(\sup_{S \in \mathcal{F}} \left| \frac{\# \text{ of flows using } S}{N} - P(S) \right| \leq \epsilon(n)\right) > 1 - \delta(n)$$

whenever

$$N \geq \max\left\{\frac{8d}{\epsilon(n)} \cdot \log \frac{13}{\epsilon(n)}, \frac{4}{\epsilon(n)} \log \frac{2}{\delta(n)}\right\}. \qquad (16)$$

Here d is the VC-dimension of $\mathcal{F}$ and $P(S)$ is the probability of a set S, which is at most $c_6\sqrt{k} \cdot \frac{r}{a}$. Thus,

$$\mathbf{Pr}\left(\sup_{S \in \mathcal{F}} \frac{\# \text{ of flows using } S}{N} \leq \frac{c_6\sqrt{k}r}{a} + \epsilon(n)\right) > 1 - \delta(n)$$

whenever condition (16) is satisfied. Let

$$\epsilon(n) = \frac{c_6\sqrt{k}r}{2a}, \text{ and } \delta(n) = \frac{2}{n}. \qquad (17)$$

Let $\frac{a}{r} = \sqrt{\frac{c \cdot n}{\log n}}$ for a constant $0 < c \leq 1/160$. Then it is sufficient that $N \geq \max\{\frac{8d\sqrt{c}}{c_6}\sqrt{\frac{n}{k \log n}} \cdot \log(\frac{26\sqrt{c}}{c_6}\sqrt{\frac{n}{k \log n}}), \frac{8\sqrt{c}}{c_6}\sqrt{\frac{n}{k \log n}} \log n\}$. When n is sufficiently large, it is sufficient that

$$N \geq \frac{4d\sqrt{c}}{c_6}\sqrt{\frac{n \log n}{k}}. \qquad (18)$$

This finishes the proof. ■

Similar to Theorem 31, we can prove the following theorem using Lemma 26.

Theorem 32: Assume that there are N random multicast sessions. There is a sequence of $\delta(n) \to 0$ such that

$$\begin{aligned}&\mathbf{Pr}\left(\forall \text{ squarelet } \mathbf{s}, \# \text{ of flows using } \mathbf{s} \geq 1\right)\\ \geq\ & 1 - \delta(n)\end{aligned}$$

when $N \geq \Omega(\max\{\log n, \frac{a^2}{kr^2}\})$.

Notice that condition (18) can always be satisfied as long as we have $n_s = \Omega(\sqrt{\frac{n \log n}{k}} \cdot \log k)$ multicast sessions. Consequently, if we assign data rate

$$\lambda = \frac{W}{2\Delta(\frac{3\sqrt{c_6}N}{2}\sqrt{k}\frac{r}{a})} = O(\frac{W}{\log n}) \qquad (19)$$

to each of the N multicast sessions (with random source node and random terminal points), with probability at least $1 - \frac{2}{n}$, the total data rate that will be routed through *every* squarelet is at most $\frac{W}{2\Delta}$. Recall that, for a multicast rooted at v_i, our routing will first send data to squarelets containing points P_i. Then for $1 \leq j \leq k-1$, we need forward the data from the squarelet containing the point $p_{i,j}$ to the nearest node $v_{i,j}$ respectively. Notice that we have proved that, for *every* squarelet, with probability at least $1 - \frac{2}{n}$, there are at least $\Theta(\log n)$ nodes inside. Thus, with probability at least $1 - \frac{2}{n}$, every data can be transferred to some node inside the squarelet. Consequently, counting this last-hop relay, with probability at least $1 - \frac{2}{n}$, the total data rate *every* squarelet has to route is at most $2\frac{W}{2\Delta} = \frac{W}{\Delta}$. Thus, these flows can be supported by a TDMA scheduling.

Consequently, we have the following theorem

Theorem 33: Assume $k \leq \theta_1 \frac{a^2}{r^2}$, there are n_s random multicast sessions and $n_s \geq \frac{4d\sqrt{c}}{c_6}\sqrt{\frac{n \log n}{k}}$. With probability at least $(1-\frac{2}{n})^2 \geq 1-\frac{4}{n}$, the achievable aggregated multicast capacity is at least

$$\Lambda_k(n) = \frac{W}{3\sqrt{c_6}\Delta\sqrt{k}\frac{r}{a}} = \Theta(\frac{W}{\sqrt{k}} \cdot \frac{a}{r}). \quad (20)$$

Theorem 34: Assume $k \leq \theta_1 \frac{a^2}{r^2}$, there are n_s random multicast sessions and $n_s \geq \frac{4d\sqrt{c}}{c_6}\sqrt{\frac{n \log n}{k}}$. With probability at least $(1-\frac{2}{n})^2 \geq 1-\frac{4}{n}$, the achievable per-flow multicast capacity is at least

$$\lambda_k(n) = \frac{W}{3\sqrt{c_6}\Delta} \cdot \frac{a}{n_s r\sqrt{k}} = \Theta(\frac{W}{n_s\sqrt{k}} \cdot \frac{a}{r}). \quad (21)$$

By setting $\frac{a}{r} = \sqrt{\frac{c \cdot n}{\log n}}$ for a constant $c \leq 1/160$, we know that the achievable aggregated multicast capacity is at least $\Theta(\sqrt{\frac{n}{k \log n}} \cdot W)$, and the per-flow multicast capacity is at least $\Theta(\sqrt{\frac{n}{k \log n}} \cdot \frac{W}{n_s})$ (which is smaller than $O(\frac{W}{\log n \cdot \log k})$, thus feasible) when there are at least $n_s = \Omega(\sqrt{\frac{n \log n}{k}} \cdot \log k)$ randomly and independently chosen flows.

C. When $k \geq \theta_1 \frac{a^2}{r^2}$

In this case, we have proved that the upper bound on the total multicast capacity is only $W/\varrho_2 = \Theta(W)$. Obviously, the total multicast capacity is at least the lower bound of the capacity for broadcast. In [9], they present a broadcast scheme to achieve capacity $\Theta(W)$. Thus, we have the following theorem

Theorem 35: The total multicast capacity $\Lambda_k(n)$ achievable by all multicast flows is at least $c_7 W$ when $k = \Omega(a^2/r^2)$, where $c_7 = \frac{1}{\Delta}$ and constant Δ is the maximum number of CDS nodes that are within interference range R of a node.

VI. Other Multicasts

A. Capacity Bound for Group Multicast

In previous sections we have studied the asymptotic multicast capacity by assuming that we randomly select $k-1$ receivers for each multicast session. In this section, we study the multicast capacity of so-called k-group multicast: for each source node v_i, there are $k-1$ groups of receivers $g_{i,1}$, $g_{i,2}$, $\cdots$, $g_{i,k-1}$. The receivers in each group $g_{i,j}$ are covered by a disk with radius $\delta \cdot r$ for a constant δ and centered at one of the receivers in the group. We assume that the center node in each group is randomly selected. The number of nodes in each group could be arbitrary. For simplicity, let node $z_{i,j}$ be the center node of group $g_{i,j}$. We then study the multicast capacity for group-multicast when each node v_i will have $k-1$ randomly selected groups and it wants to send data with rate λ_i to all receivers in these $k-1$ groups.

As the case when each group has only one node, when $k \geq \theta_1 a^2/r^2$, it is easy to prove that the capacity for group-multicast is at most W/ϱ_2 as Theorem 19. Clearly, a simple broadcast based on the connecting dominating set constructed previously will also achieve a capacity for group-multicast at least $\frac{W}{\Delta}$. Consequently, we have

Theorem 36: For group-multicast, when $k \geq \theta_1 a^2/r^2$ for any constant $\theta_1 > 0$, the capacity of group-multicast is at most W/ϱ_2 and at least $\frac{W}{\Delta}$.

First, for group-multicast, a multicast tree has to reach the center node $z_{i,j}$ of each group $g_{i,j}$. Then from Theorem 15, the capacity of group-multicast is at most $c_1 \cdot \frac{aW}{r\sqrt{k}}$ with high probability when $k < \theta_1 \cdot a^2/r^2$. We then show how to design multicast routing for the group-multicast problem: we first apply our multicast scheme for traditional multicast when nodes $z_{i,j}$, $1 \leq j \leq k-1$, are receivers for source node v_i. We then let node $z_{i,j}$ multicast locally to all receivers in the group $g_{i,j}$. We already proved in Theorem 21, the total length of the multicast tree to span these randomly selected nodes $z_{i,j}$ is at most $c_3\|EMST(U_i')\|$ with high probability. Recall that $\|EMST(U_i')\| \leq \frac{3\tau(2)\sqrt{k}\cdot a}{2}$ with high probability. Notice that the total area covered by transmitting disks of relay nodes used for relaying data from each $z_{i,j}$ to receivers in its group is at most $\pi(\delta+1)r^2$. Then, the area covered by all transmitting disks for a multicast session is at most

$$\begin{aligned} & 2r \cdot |MT(P_i)| + (k-1) \cdot \pi(\delta+1)r^2 \\ \leq\; & 2r \cdot c_3 \frac{3\tau(2)\sqrt{k} \cdot a}{2} + k \cdot \pi(\delta+1)r^2 \quad (w.h.p.) \\ \leq\; & \left(\theta_1 \pi(\delta+1)^2 + 3c_3\tau(2)\right) \cdot \sqrt{k} \cdot a \cdot r \end{aligned}$$

The last inequality comes from the fact that $k \leq \theta_1 \frac{a^2}{r^2}$. For convenience, let $c_8 = \theta_1\pi(\delta+1)^2 + 3c_3\tau(2)$. Then similar to Theorem 23, we can prove that the number of nodes that will get a copy of the data from one multicast session is at most $c_8\sqrt{k} \cdot a \cdot r \cdot \frac{2n}{a^2} = 2c_8 \cdot n \cdot \sqrt{k} \cdot \frac{r}{a}$. Thus, we have

Theorem 37: When $k \leq \theta_1 a^2/r^2$, the aggregated multicast capacity for group-multicast with $k-1$ groups is at most $c_1 \cdot \frac{aW}{r\sqrt{k}}$, and is at least $c_9 \frac{a \cdot W}{r\sqrt{k}}$, with high probability. Here constant $c_9 = 2c_8\Delta$.

B. Bounds for Arbitrary Networks

In previous studies we concentrated on the multicast capacity for random networks when nodes will be randomly placed in the deployment region. In this section we will study what is the asymptotic maximum multicast capacity that can be achieved by a specific *connected* network when nodes' position can be carefully selected.

We first present a constructive lower bound on the multicast capacity. Assume that n nodes are deployed in a $\sqrt{n}$ by $\sqrt{n}$ grid, each cell has side-length r, *i.e.* the side-length of the square is $a = r\sqrt{n}$. The $k-1$ receivers are randomly selected from the grid points. We then perform multicast as before: the multicast tree is constructed based on the Euclidean minimum spanning tree connecting source node and $k-1$ receivers. Let L be the total length of the Euclidean MST constructed above. Similar to previous studies in Subsection V-B, we know that the multicast capacity Λ_n satisfies $\Lambda_n \geq c_5' \frac{W \cdot a^2}{L \cdot r}$ for some constant c_5' depending only on R/r. Lemma 12 gives an upper bound on L for the Euclidean minimum spanning tree. This implies the following corollary:

Corollary 38: The multicast capacity Λ_n of n nodes for an arbitrary network, where we can choose node positions, is at least $\frac{c_5'}{2\sqrt{2}} \frac{1}{\sqrt{k}} \frac{a}{r} \cdot W$.

We then present an upper-bound on the multicast capacity of an arbitrary network of n nodes deployed in a square of side-length a meters, where n is sufficiently large. Consider a bit b that is sent from the source node to all receivers. Let A be the average area a bit has to cover. Notice that when a node transmit a bit b, then all nodes within its transmission range cannot receive data from other nodes due to interference, thus, we say the bit b covers the transmission disk of the transmitting node. Obviously, we have

$$\Lambda_n \cdot A \leq W \cdot a^2,$$

where Λ_n is the aggregated multicast capacity of all nodes. Left-side of the inequality denotes the total bits$\times$meter^2 achieved by all multicast sessions, while the right-side of the inequality is derived from the fact that, at any time, the simultaneously transmitting disks must be disjoint.

C. Capacity Bounds for d-dimensional Networks

It is not difficult to extend our capacity bounds to networks in d dimensions. Assume that the network nodes are randomly deployed in a d-dimensional cube with side-length a and every node has a transmission range r and an interference range R. The total Euclidean length of edges in any tree spanning k nodes randomly distributed in the d-dimensional cube is, with high probability, at least $\Omega(k^{\frac{d-1}{d}} a)$ [17]. Similar to 12, we can show that the total Euclidean length of the edges in EMST spanning any k nodes is at most $O(k^{\frac{d-1}{d}} a)$. The total volume of the region shaded by all transmitting spheres of all internal nodes of a multicast tree is then at the order of $\Theta(k^{\frac{d-1}{d}} \cdot \frac{a}{r} \cdot r^d)$. Consequently, the number of nodes that will get a copy of the multicast data is at the order of $\Theta(k^{\frac{d-1}{d}} \cdot \frac{a}{r} \cdot r^d \cdot \frac{n}{a^d})$, which is at most n when $k \leq \theta_d(\frac{a}{r})^d$ for some constant θ_d depending only on dimension d. This implies that the aggregated multicast capacity of n_s multicast sessions is at the order of $\Theta(\frac{nW}{nk^{\frac{d-1}{d}}(\frac{r}{a})^{d-1}}) = \Theta(\frac{W}{k^{\frac{d-1}{d}}} \cdot (\frac{a}{r})^{d-1})$. To show the lower-bound on achievable multicast capacity, we will also partition the cube into cubelets with side length $r/\sqrt{3+d}$ such that two nodes inside two cubelets sharing a $(d-1)$-dimensional facet can communicate with each other directly. We can then show that, given a specific cubelet, the probability that a random multicast flow will be routed through this cubelet is $k^{\frac{d-1}{d}} \cdot (\frac{r}{a})^{d-1}$. Consequently, we have the following theorem:

Theorem 39: Assume that n nodes are randomly placed in a d-dimensional cube of side-length a and each node has a fixed transmission range r. The aggregated multicast capacity of n_s random multicast sessions (each with k terminals) is

$$\Lambda_k(n) = \begin{cases} \Theta((\frac{a}{r})^{d-1} \cdot \frac{W}{k^{\frac{d-1}{d}}}) & \text{when } k \leq \theta_d(\frac{a}{r})^d, \text{ and} \\ & n_s = \Omega((\frac{a}{r})^{d-1} \frac{\log n}{k^{\frac{d-1}{d}}}). \\ \Theta(W) & \text{when } k \geq \theta_d(\frac{a}{r})^d \end{cases} \tag{22}$$

Recall that Penrose [14] showed that the asymptotic length of the longest edge of EMST of n nodes randomly placed in a d-dimensional cube of unit side-length is $(\frac{2(1-1/d)\log n}{\theta n})^{1/d}$ where θ is the volume of the unit ball in d-dimension. Thus, to get a connected network, we need a and r satisfy the following condition, with high probability,

$$\frac{a}{r} \leq c_d \sqrt[d]{\frac{n}{\log n}}$$

for some constant c_d depending on dimension d. Thus, we have the following theorem.

Theorem 40: If we can choose the transmission range r or the deployment region a, the aggregated multicast capacity of n_s multicast sessions for n randomly placed nodes in a d-dimensional cube of side-length a is

$$\Lambda_k(n) = \begin{cases} \Theta(W \cdot (\frac{n}{k \log n})^{1-\frac{1}{d}}) & \text{when } k \leq \theta_d(c_d)^d \frac{n}{\log n}, \text{ and} \\ & n_s = \Omega((\frac{n}{k \log n})^{1-\frac{1}{d}} \cdot \log n). \\ \Theta(W) & \text{when } k \geq \theta_d(c_d)^d \frac{n}{\log n} \end{cases} \tag{23}$$

VII. Literature Reviews

Network capacity has been extensively studied recently. Capacity can be generalized to the notion of *capacity region* for fixed networks. For a given statistical description of the network, a set of constraints (such as power per node, link capacity, *etc.*), and a list of desired communication pairs, the capacity region is the closure of all rate tuples that can be achieved simultaneously. Here a rate tuple specifies the rate for each of the desired communications. Kyasanur and Vaidya [12] studied the capacity region on random multi-hop multi-radio multi-channel wireless networks when there are total c channels available and each node has m wireless interfaces with $m \leq c$. On the other aspect, several papers [2], [11] recently studied how to satisfy a certain traffic demand vector from all wireless nodes by a joint routing, link scheduling, and channel assignment under certain wireless interference models.

Gupta and Kumar [6] studied the asymptotic *unicast* capacity of a multi-hop wireless networks for two different models. When each wireless node is capable of transmitting at W bits per second using a constant transmission range, the throughput obtainable by *each* node for a randomly chosen destination is $\Theta(\frac{W}{\sqrt{n \log n}})$ bits per second under a non-interference protocol, where n in number of nodes. If nodes are optimally assigned and transmission range is optimally chosen, even under optimal circumstances, the throughput is only $\Theta(\frac{W}{\sqrt{n}})$ bits per second for each node. Similar results also hold for physical interference model. Notice that the results presented in [6] did not consider the additional burden in coordinating access to wireless channels, the effect of mobility and link failures, the effect of the need to route traffic in a distributed way. They also did not address the delay of the route. The delay could caused by burst traffic or when nodes are mobile and links are not stable. It can also be imagined that using directional antennas or beam-forming will help to improve the spatially concurrency of transmissions and thus the capacity of the networks.

Grossglauser and Tse [5] recently showed that mobility actually can help to improve the unicast capacity if we allow arbitrary large delay. Their main result shows that the average long-term throughput per source-destination pair can be kept constant even as the number of nodes per unit area increases. Notice that this is in sharp contrast to the fixed network scenario (when nodes are static after random deployment). The main idea used in [5] is to use some intermediate node to serve as ferry node: this node will carry the data from the source node and move around and it will dump the data to the target node when it is within its communication range. In other words, essentially, the result presented in [5] still obey the capacity bound proposed in [6]: the capacity is improved because the average distance $\overline{L}$ a packet has to be transmitted is reduced from $\Theta(1)$ in [6] to $\Theta(r(n))$ in [5]. In summary, for random networks, under the protocol model, the achievable per-flow throughput capacity $\lambda(n)$ and the average travel distance $\overline{L}$ satisfies $\lambda(n) \cdot \overline{L} \leq \Theta(\frac{W}{\Delta^2 n \cdot r(n)})$. Similar phenomenon has also been observed in [13]. They found that the traffic pattern determines whether the per flow capacity of a wireless network will scale to large networks. They observed that non-local traffic patterns in which the average distance grows with the network size result in a rapid decrease of per flow capacity. They also examined the interactions of the 802.11 MAC and the ad hoc forwarding and the effect on the capacity of wireless networks. Although 802.11 discovers reasonably good schedules, they nonetheless observed capacities markedly less than the optimal even for very simple networks, such as chain and lattice networks, with very regular traffic patterns. This confirms the importance of using carefully designed transmission schedule to improve the network throughput whenever it is possible.

Broadcast capacity of an arbitrary network has been studied in [9], [18]. They essentially show that the broadcast capacity of a given network is $\Theta(W)$ for single source broadcast and the achievable broadcast capacity per flow is only $\Theta(W/n)$ if each of the n nodes will serve as source node. The upper bound $\Theta(W)$ on broadcast capacity trivially holds since each node can receive at most W bits/sec. The capacity $\Theta(W)$ is achieved by constructing a connected dominating set in which we can schedule every node in CDS to transmit at least once in constant time slots. This capacity bounds also apply to random networks. Keshavarz-Haddad *et al.* [10] studied the broadcast capacity with dynamic power adjustment for physical interference model.

Multicast capacity was not fully studied in the literature. Jacquet and Rodolakis [8] studied the scaling properties of multicast for random wireless networks. They essentially studied the *normalized multicast cost*, which is defined as the ratio of the number of links in the multicast tree over the average route length from a random source in the multicast group to a random destination in the multicast group. They briefly showed that the maximum rate at which a node can transmit multicast data is $O(\frac{W}{\sqrt{kn\log n}})$. At the same time as our results, Shakkottai *et al.* [16] studied the multicast capacity of random networks when the number of multicast sources is n^{ϵ} for some $\epsilon > 0$, and the number of receivers per multicast flow is $n^{1-\epsilon}$. They assume the protocol interference model and use the dense random network model. They show that the sum of the source rates $\Lambda(n)$ that the network can support is $O(\frac{\sqrt{n^{\epsilon}}}{\sqrt{\log n}})$ *w.h.p.*, with a per flow throughput capacity of $O(\frac{1}{\sqrt{n^{\epsilon}\log n}})$ *w.h.p.*. Notice that this result can be implied by our results using $n_s = n^{\epsilon}$ and $k = n^{1-\epsilon}$. To achieve the upper bound, they propose a simple and novel routing architecture, called the *multicast comb*, to transfer multicast data in the network.

VIII. Conclusions

In this paper, we essentially studied the multicast capacity that can be achieved by some random wireless networks. We derive analytical upper bounds and lower bounds on multicast capacity of a wireless network when all nodes are uniformly and randomly deployed in a square region with side-length a, and all nodes have the same transmission range r. We show that the total multicast capacity is only $\Theta(\sqrt{\frac{n}{\log n}} \cdot \frac{W}{\sqrt{k}})$ when $k = O(\frac{n}{\log n})$; the total multicast capacity is $\Theta(W)$ when $k = \Omega(\frac{n}{\log n})$. We also studied the multicast capacity for group-multicast and for arbitrary networks.

Observe that all our results are proved when the deployment region is a square with side-length a and the transmission range of all nodes is uniform with value r. It is not difficult to show that all our results still apply when the deployment region is a square with side length $a = 1$, while the transmission range is selected appropriately, *i.e.*, $r = \Theta(\sqrt{\frac{\log n}{\pi n}})$. It is also not difficult to show that our results still hold when $r = 1$ while the deployment region has a bounded aspect ratio such as a disk. Further, we considered the protocol interference model for random networks. It is not difficult to show that our results still hold (with different constants) when we apply the physical interference model (where all nodes have *fixed* uniform transmission power P) and the signal power at distance d decays as $\frac{1}{d^{\alpha}}$ for $\alpha > 2$. The basic idea is to show that, for such physical interference model, there is a *logic* transmission range r and interference range R (with $R = \Theta(r)$) such that when $\|u - v\| \leq r$ and *no* other transmitting nodes within distance R of receiving node v, node u can always successfully send data to v. All computations will be similar by using such logic transmission range and interference range. The details of all computations are omitted here due to space limit.

There are some interesting questions left for study for multicast capacity. The first question is what is the multicast capacity when the link capacity is not uniform: shorter links will have larger capacity. The second question is what is the multicast capacity when the Gaussian channel is used, instead of assuming that each node has a constant transmission range and has a constant data rate W. Last but not the least question is what is the tradeoffs between the delay and multicast capacity for random mobile networks?

References

[1] G.-S. Ahn, E. Miluzzo, A. T. Campbell, S. G. Hong, and F. Cuomo. Funneling-mac: A localized, sink-oriented mac for boosting fidelity in sensor networks. In *In Proceedings of the First ACM Conference on Embedded Networked Sensor Systems (SenSys)*, 2006.

[2] M. Alicherry, R. Bhatia, and L. E. Li. Joint channel assignment and routing for throughput optimization in multi-radio wireless mesh networks. In *MobiCom '05: Proceedings of the 11th annual international conference on Mobile computing and networking*, pages 58–72, New York, NY, USA, 2005. ACM Press.

[3] K. Alzoubi, X.-Y. Li, Y. Wang, P.-J. Wan, and O. Frieder. Geometric spanners for wireless ad hoc networks. *IEEE Transactions on Parallel and Distributed Processing*, 14(4):408–421, 2003. Short version in IEEE ICDCS 2002.

[4] D.-Z. Du and F.-K. Hwang. A proof of the gilbert-pollak conjecture on the steiner ratio. *Algorithmica*, 7((2,3)):121–135, 1992.

[5] M. Grossglauser and D. Tse. Mobility increases the capacity of ad-hoc wireless networks. In *INFOCOMM*, volume 3, pages 1360 –1369, 2001.

[6] P. Gupta and P. Kumar. Capacity of wireless networks. Technical report, University of Illinois, Urbana-Champaign, 1999.

[7] P. Gupta and P. R. Kumar. Critical power for asymptotic connectivity in wireless networks. *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming, W. M. McEneaney, G. Yin, and Q. Zhang (Eds.)*, 1998.

[8] P. Jacquet and G. Rodolakis. Multicast scaling properties in massively dense ad hoc networks. In *ICPADS '05: Proceedings of the 11th International Conference on Parallel and Distributed Systems - Workshops (ICPADS'05)*, pages 93–99, Washington, DC, USA, 2005. IEEE Computer Society.

[9] A. Keshavarz-Haddad, V. Ribeiro, and R. Riedi. Broadcast capacity in multihop wireless networks. In *MobiCom '06: Proceedings of the 12th annual international conference on Mobile computing and networking*, pages 239–250, New York, NY, USA, 2006. ACM Press.

[10] A. Keshavarz-Haddad and R. Riedi. On the broadcast capacity of multihop wireless networks: Interplay of power, density and interference. In *4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, 2007.

[11] M. Kodialam and T. Nandagopal. Characterizing achievable rates in multi-hop wireless networks: the joint routing and scheduling problem. In *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 42–54, New York, NY, USA, 2003. ACM Press.

[12] P. Kyasanur and N. H. Vaidya. Capacity of multi-channel wireless networks: impact of number of channels and interfaces. In *MobiCom '05: Proceedings of the 11th annual international conference on Mobile computing and networking*, pages 43–57, New York, NY, USA, 2005. ACM Press.

[13] J. Li, C. Blake, D. S. J. D. Couto, H. I. Lee, and R. Morris. Capacity of ad hoc wireless networks. In *ACM MobiCom*, 2001.

[14] M. Penrose. The longest edge of the random minimal spanning tree. *Annals of Applied Probability*, 7:340–361, 1997.

[15] P. Santi and D. M. Blough. The critical transmitting range for connectivity in sparse wireless ad hoc networks. *IEEE Trans. on Mobile Computing*, 2:25–39, Mar. 2003.

[16] S. Shakkottai, X. Liu, and R. Srikant. The multicast capacity of ad hoc networks. In *Proc. ACM Mobihoc*, 2007.

[17] J. M. Steele. Growth rates of euclidean minimal spanning trees with power weighted edges. *The Annals of Probability*, 16(4):1767–1787, Oct 1988.

[18] B. Tavli. Broadcast capacity of wireless networks. *IEEE Communication Letters*, 10(2), February 2006.

[19] V. Vapnik and A. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, 16(2):264–280, 1971.

[20] P.-J. Wan, K. M. Alzoubi, and O. Frieder. Distributed construction of connected dominating set in wireless ad hoc networks. In *INFOCOM*, 2002.

[21] F. Xue and P. R. Kumar. The number of neighbors needed for connectivity of wireless networks.

[22] R. Zheng. Information dissemination in power-constrained wireless networks. In *Proc. of The 25th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 2006.