Impact of Deployment Size on the Asymptotic Capacity for Wireless Ad Hoc Networks Under Gaussian Channel model

Shao-Jie Tang[†], XuFei Mao[‡], Xiang-Yang Li[*†], and Cheng Wang[§]
[*] Institute of Computer Application Technology, Hangzhou Dianzi University, Hangzhou, PRC.
[†] Department of Computer Science, Illinois Institute of Technology, Chicago, IL, USA.
[‡] Department of Computer Science, BeiJing University of Post and Telecommunications, PRC.
[§] Department of Computer Science, TongJi University, PRC.
{stang7,xmao3}@iit.edu, xli@cs.iit.edu

***Abstract*—We study the *throughput capacity* and *transport capacity* for both random and arbitrary wireless networks under *Gaussian Channel* model when all wireless nodes have the same constant transmission power P and the transmission rate is determined by Signal to Interference plus Noise Ratio (SINR). We consider networks with n wireless nodes $\{v_1, v_2, \cdots, v_n\}$ (randomly or arbitrarily) distributed in a square region B_a with a side-length a. We randomly choose n_s node as the source nodes of n_s multicast sessions. For each source node v_i, we randomly select k points and the closest k nodes to these points as destination nodes of this multicast session. We derive achievable lower bounds and some upper bounds on both throughput capacity and transport capacity for both unicast sessions and multicast sessions. We found that the asymptotic capacity depends on the size a of the deployment region, and it often has three regimes.**

***Index Terms*—Wireless networks, throughput capacity, transport capacity, unicast, multicast, Gaussian channel.**

I. Introduction

Recently, the network capacity has been studied extensively under different network, system, and interference models. The ground breaking work of Gupta *et al.* [3] has shown that when n wireless nodes are randomly placed in a square region with side-length 1, for randomly picked n pairs of source/destination nodes, the total information exchangeable by each pair will go to zero in order of $\frac{1}{\sqrt{n \log n}}$ as n tends to ∞ under the protocol interference model (PrIM) and physical interference model. They also showed in [3] that even all nodes are located optimally, the amount of information that can be exchanged by each source/destination pair still goes to zero in order of $\frac{1}{\sqrt{n}}$. In addition, the authors of [1], [2] proposed alternative techniques that achieve unicast capacity $\Theta(\frac{1}{\sqrt{n \log n}})$ for random wireless networks. Recently, Francheschetti *et al.* [4] proved that per-flow unicast capacity of order $\frac{1}{\sqrt{n}}$ is also achievable in networks of randomly located nodes when Gaussian channel model is used. Hence, the unicast capacity gap between randomly and arbitrary wireless networks is claimed to be closed. However, the work in [4] is based on Gaussian channel model, while the work in [3] is based on PrIM and physical interference models, and the asymptotic capacity bounds may be different in different models.

Observe that the majority researches in the literature assumed that n nodes are deployed in a 2 dimensional square of side length 1 or $\sqrt{n}$. Assuming that n nodes are randomly placed in a square region of side-length a, the main purpose of this paper is to study the impact of the deployment region size a on the asymptotic *unicast capacity* and more generally the multicast capacity, of large scale random or arbitrary wireless networks under Gaussian channel model. Two different capacities will be studied, namely, the throughput capacity and the transport capacity. Under Gaussian channel model, the data rate between any pair of transceivers (u, v) is determined by several parameters, including transmission power P of u, the noise N_0, the interference signals from all other simultaneously transmitting nodes rather than u. Hence, pairs of nodes can communicate directly with different data rates. For presentation simplicity, we assume that there is only one channel in wireless networks. And as always, data are sent from node to node either by one-hop or by multi-hop manner until they reach the destination. In addition, we assume every node has enough buffer to save the relay traffic temporarily while waiting for being transmitted such that no packet will be lost through relaying.

Our Main Contributions: In this work we derive analytical upper bounds and lower bounds of unicast(multicast) capacity for wireless networks under Gaussian channel model when n wireless nodes (randomly or arbitrarily) distributed in a square region with side-length a. We studied different cases when a is in different range, *i.e.*, a is a function of n. In general, we observe that the capacity of a random (or arbitrary) network has three different regimes depending on the deployment size a. Building on existing milestone results, our main contributions are to firstly present the tight bounds on the impact of the deployment size a on the capacity for majority cases. Our main results are summarized as follows:

Theorem 1: For an arbitrary wireless network, the total

The research of authors are partially supported by NSF CNS-0832120, NSF CNS-1035894, National Natural Science Foundation of China under Grant No. 60828003, National 973 Project of China under Grant No.2011CB302700, program for Zhejiang Provincial Key Innovative Research Team, program for Zhejiang Provincial Overseas High-Level Talents (One-hundred Talents Program), and Tsinghua National Laboratory for Information Science and Technology(TNList).

unicast throughput capacity $\Lambda(n)$ is:

$$\Lambda(n)=\begin{cases}\Theta(1) & \text{when } a=O(1)\\ \Theta(a^2) & \text{when } a=o(\sqrt{n})\\ \Theta(n) & \text{when } a=\Omega(\sqrt{n})\end{cases} \tag{1}$$

Theorem 2: For a random wireless network, the minimum per-flow unicast throughput capacity is:

$$\varphi(n)=\begin{cases}\Theta(\frac{1}{n}) & \text{if } a=O(1)\\ \Theta(\frac{a}{n}) & \text{if } a=o(\sqrt{n})\\ \Theta((\frac{a}{\sqrt{n}})^{-\beta}\cdot\frac{1}{\sqrt{n}}) & \text{when } a=\Omega(\sqrt{n})\end{cases} \tag{2}$$

Theorem 3: For a random wireless network, an upper bound (partially achievable) of minimum per-flow multicast throughput capacity, when $n_s=\Theta(n)$, is:

$$\varphi_k(n)=\begin{cases}O(\frac{a}{n\sqrt{k}}) & \text{if } a=O(\sqrt{n})\\ \Theta(\frac{1}{\sqrt{n}}\cdot\frac{1}{\sqrt{k}}) & \text{if } a=\Theta(\sqrt{n}), k=O(\frac{n}{\lg^2 n})\\ O((\frac{a}{\sqrt{n}})^{-\beta}\frac{1}{\sqrt{n}}\cdot\frac{1}{\sqrt{k}}) & \text{if } a=\Omega(\sqrt{n}), k=O(\frac{n}{\lg^2 n})\end{cases} \tag{3}$$

Theorem 4: Consider a random wireless networks, where nodes following a Poisson distribution with parameter $\frac{n}{a^2}$ are distributed in B_a with $a=\Omega(\sqrt{n})$. Assume that n_s random multicast flows are generated. Under Gaussian channel model, the per-flow multicast capacity $\varphi_k(n)$ is at most

$$\varphi_k(n)=\begin{cases}O((\frac{a}{\sqrt{n}})^{-\beta}\cdot\frac{1}{n_s}\cdot\frac{\sqrt{n}}{\sqrt{k}}) & \text{if } k\le\frac{n}{(\log n)^{\beta}}\\ O(\frac{n}{n_s k}(\frac{a}{\sqrt{n}})^{-\beta}(\log n)^{-\frac{\beta}{2}}), & \text{if } k\ge\frac{n}{(\log n)^{\beta}}\end{cases} \tag{4}$$

In contrast to [3], [4], studying unicast capacity of wireless network under Gaussian channel model needs new technical insight. One of reasons is that the interference concept under Gaussian channel model is different from PrIM. Under PrIM, every node has fixed transmission range and interference range, the data rate between them is fixed as well, thus a necessary condition for any two nodes to be able to communicate with each other is that 1) these two nodes must be in the transmission range of each other; and 2) when a transmitter is active, all other nodes within its interference range cannot transmit simultaneously. Compared with PrIM, Gaussian channel model gives better link rate when other transmissions are treated as noise. The data rate under Gaussian channel model is determined by power, distance and noise. Any two nodes can communicate with each other although the data rate maybe go to zero when the distance between the transceiver pair is long or there are too much noise. Hence, some techniques used in previous work cannot be applied directly to Gaussian channel model without modification.

The rest of the paper is organized as follows. In Section II we discuss in detail the network model used. We present both upper-bounds and lower-bounds of unicast capacity for an arbitrary wireless network in Section III. The unicast capacity bounds for random wireless networks are presented in Section IV. We study the multicast capacity in Section V. We review the related results in Section VI and conclude the paper in Section VII.

II. System Model and Preliminaries

A. Network Model and Asymptotic Capacity

Consider a square region B_a with side length a. We assume that there is a set $V=\{v_1,v_2,\cdots,v_n,\cdots\}$ *ordinary* wireless terminals deployed in B_a following Poisson distribution with parameter $\frac{n}{a^2}$. In other words, given a region X with area x, the probability that there are exactly k nodes inside X is $\frac{(n/a^2)^k e^{-xn/a^2}}{k!}$. The expected number of nodes located in the region B_a is n. We randomly pick n_s out of n wireless terminals as source nodes. Here, n_s can be as large as n which means that every node will serve as a source node. For each source node v_i, we randomly select a point p_i in B_a and the node which is closest to p_i will become the destination node of v_i for unicast. Here, if the source node v_i chooses itself as its destination node, we can randomly generate point p_i again to avoid this. For studying multicast capacity, we assume that each multicast session will have k receivers. For each source node v_i, we randomly pick k points $p_{i,j}$, $1\le j\le k$, in B_a and then the closest node $v_{i,j}$ to $p_{i,j}$ will serve as a destination node of the ith flow that has the source node v_i.

We assume that all nodes have a constant transmission power P, and for each transceiver pair (v_i,v_j), node v_j receives the transmitted signal from node v_i with power $P\cdot\ell(d(v_i,v_j))$, where $d(v_i,v_j)$ is the Euclidean distance between node v_i and v_j, $\ell(d)$ is the transmission loss during a path with length d. In this paper, we consider the attenuation function

$$\ell(d)=\min\{1,d^{-\beta}\}$$

where the constant $\beta>2$. Hence, any two nodes can establish a direct communication link over a unit bandwidth channel, of rate $R(v_i,v_j)=$

$$\log(1+\frac{S(v_i,v_j)}{N_0+I(v_i,v_j)})=\log(1+\frac{P\cdot\ell(v_i,v_j)}{N_0+\sum_{v_q\in T(i)}P\cdot\ell(v_q,v_j)})$$

Here, $T(i)$ is set of nodes transmitting simultaneously with v_i and N_0 is the variance of background noise, usually be a constant, $I(v_i,v_j)$ is the total interference at the receiving node v_j when v_i and v_j communicate, and $S(v_i,v_j)$ denotes the strength of signal received by v_j sent from v_i.

Capacity Definition: In this work we will study both the asymptotic throughput capacity and asymptotic transport capacity. We assume that any node v_i could serve as the source node for some unicast or multicast. And for each source node v_i, assume that node v_i will send data to its receiver(s) by unicast (or multicast) with a data rate λ_i.

Let $\lambda=(\lambda_1,\lambda_2,\cdots,\lambda_{n-1},\lambda_n)$ be the *rate vector* of the multicast data rate of all multicast sessions. 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 unicast (multicast) data rate λ_i for each source node v_i are all fixed, a multicast rate vector λ 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 destination nodes in set U_i. 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 can send $T\cdot\lambda_i$ bits to its corresponding receiver(s).

Definition 1 (Throughput Capacity): Given a feasible rate vector λ, the *total throughput* of λ is $\Lambda_k(n) = \sum_{i=1}^{n}\lambda_i$. The *average throughput* is $\lambda_k(n) = \frac{\sum_{i=1}^{n}\lambda_i}{n_s}$, where n_s is the number of unicast (multicast) sessions, and k is the total number of nodes in each unicast(multicast) session, including the source node. Given n_s sessions with S as source nodes, the *minimum per-flow throughput capacity* is defined as

$$\varphi_k(n) = \min_{v_i\in S}\lambda_i. \tag{5}$$

In this work, we will focus on the minimum per-flow throughput capacity. Given any set of successful transmissions taking place over time and space, we define the transport capacity as the ability for the network to transmit bits to their destinations with a distance of one meter (or unit) per second.

Definition 2 (Transport Capacity): An aggregated unicast *transport capacity* $\Lambda_k^T(n)$ bit-meters/sec is defined as

$$\Lambda_k^t(n) = \sum_{i=1}^{n} l_i\lambda_i \tag{6}$$

Here, l_i is the length of the path connecting source node s_i and destination node t_i in i^{th} unicast session, and λ_i is the feasible (achievable) data rate between s_i and t_i.
From now on, we use $\lambda^T(u,v)$ to denote the transport capacity between a pair of nodes u, v.

Definition 3 (Capacity of Random Networks): For a class of random networks, the per-session asymptotic throughput capacity is $\Theta(g(n))$ if there are constants $0 < c < c'$, s.t.,

$$\lim_{n\to+\infty}\Pr(\varphi_k(n) = cg(n) \text{ is feasible}) = 1,$$
$$\liminf_{n\to+\infty}\Pr(\varphi_k(n) = c'g(n) \text{ is feasible}) < 1.$$

NOTATIONS: Throughput this paper, for a continuous region B_a, we use $|B_a|$ 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.

B. Technical Lemmas

To study the asymptotic capacity, we first present several technique lemmas that are essential for the analysis. For a random wireless network with n wireless nodes located in a square region B_a with side length a, we partition B_a into cells with side length c. Two nodes u and v are said to have *cell-distance* d if they are located in two cells that are separated by d-cells. Please refer to Appendix for corresponding proof.

Lemma 5: Based on a TDMA schedule, for any transceiver pair (u,v) with cell-distance d, the data rate $R(u,v)$ only depends on d and c. Furthermore, when $c\cdot d\to\infty$, $R(u,v) = \Omega(c^{-\beta}d^{-\beta-2})$.

Lemma 6: If we partition the square B_a with side length a into $\frac{a^2}{c^2}$ cells with constant side length c, then *w.h.p.*, there are less than $\log\frac{a}{c}\times\frac{nc^2}{a^2}$ nodes in each cell.

Lemma 7: If we partition square B_a into stripes with width a and height c_1, then with probability at least $1 - \frac{\sqrt{n}}{c_1}e^{-c_1\sqrt{n}}(\frac{e}{2})^{2c_1\sqrt{n}}$ the number of nodes in each stripe will be no more than $2\frac{c_1}{a}\cdot n$.

C. Highway System and related

Some of our routing strategies are built upon the highway system developed in [4]. Here we briefly review its construction and some key properties. To begin the construction of highway system, we partition the deployment box B_a into $m = \frac{a}{\sqrt{2}c}$cells with a side length c. By appropriately choosing c, we can arrange that the probability that a square contains at least a Poisson point is as high as we want. Here when $a = O(\sqrt{n})$, choosing c as some constant is enough, while when $a = \Omega(\sqrt{n})$, we choose $c = \theta_1\cdot\frac{a}{\sqrt{n}}$ for some constant θ_1.

Then based on percolation theorem, we can choose c large enough such that with high probability (w.h.p.) there are paths crossing B_a from left to right. These paths can be grouped into disjoint sets of paths: each group has $\lceil\delta\log m\rceil$ paths, crossing a rectangle of width m and height $(\kappa\log m + \epsilon_m)$ cells, for all $k > 0$, δ small enough, and a vanishingly small ϵ_m so that the side length of each rectangle is an integer. Same results still hold when looking for paths crossing B_a from bottom to top. Then by union bound, we claim that there exist both horizontal and vertical disjoint paths *w.h.p.*. These paths are called the *highway system*. From now on, we simply call a node *highway node* if the node is on one of horizontal or vertical (or both) paths, otherwise, it will be called *non-highway node*.

Then we slice the network area into horizontal strips of constant width c_0 such that there are at least as many paths as slices inside each rectangle of size $m\times\kappa\log m + \epsilon_m$ by choosing c_0 appropriately. Then we impose that nodes from the ith slice communicate directly with the ith horizontal path. And it is also proved in [4] that *w.h.p.*, there are at most $\Theta(\sqrt{n})$ nodes contained in each stripe. Finally, we can get the following important lemma.

Lemma 8: [4] The nodes along the highways can achieve *w.h.p.*, a per-flow rate of $\Omega(\frac{1}{\sqrt{n}})$.

III. UNICAST CAPACITY FOR ARBITRARY NETWORKS

We first study unicast throughput and transport capacity for an arbitrary wireless network where n nodes arbitrarily distributed in a square region with side length a.

A. When $a = O(1)$

Lemma 9: For an arbitrarily network, when side length of square $a = O(1)$, the total unicast throughput capacity for n_s transceiver pairs is $\Theta(1)$. In addition, the transport capacity is $\Theta(a)$ as well.

Proof: The lower bound is clearly $\Omega(1)$ since, in any time slot we pick only one transceiver pair (u and v) to communicate, all other transmitters are silent. In this case, the rate $R(u,v) = \log(1 + \frac{P\cdot\ell(u,v)}{N_0+I(u,v)}) = \Omega(1)$ because the Euclidean distance $\ell(u,v)$ is at most $\sqrt{2}a$ which is a constant, $I(u,v)$ is zero in this case and N_0 is a constant. Thus, the lower bound of unicast throughput capacity for n unicast sessions is $\Omega(1)$. Clearly, when we pick two nodes within distance $\Theta(a)$, the transport capacity can achieve $\Omega(1\times a)$.

We then show that the capacity is $O(1)$ by the following observations. Assume for any time slot t, there are $m\geq 2$ simultaneously active links in the network. Then for any transceiver pair u and v, the rate $R(u,v)=\log(1+\frac{P\cdot\ell(d(u,v))}{N_0+I(u,v)})\le\log(1+\frac{P\cdot 1}{N_0+\sum_{m-1}P\cdot 1})=\log(1+\frac{1}{\frac{N_0}{P}+(m-1)})\le\frac{1}{\frac{N_0}{P}+(m-1)}\le\frac{1}{m-1}=O(\frac{1}{m})$. The last inequality is true since $\log(1+x)\le x$). Since there are m simultaneously active links for time slot t, the total throughput capacity of the network is $O(1)$. Clearly, in this case, the upper bound on transport capacity is $\frac{m}{m-1}\cdot a=O(a)$ bits-meters/sec. This completes the proof. ■

Obviously, the per flow network throughput capacity is $\Theta(\frac{1}{n})$.

B. When $a=\Omega(1)$, and $a=O(\sqrt{n})$

Lemma 10: For any pair of source/destination nodes (u,v), the transport capacity of a direct link $e=(u,v)$ is $\log(1+\frac{P\cdot\ell(d)}{N_0+I(u,v)})\cdot d$ where d is the Euclidean distance between node u and node v. In addition, the transport capacity between u and v will get its maximum value when $d=1$.

Proof: According to the definition, the transport capacity between link (u,v) is

$$\lambda^T(u,v)=\log(1+\frac{P\cdot\ell(d)}{N_0+I(u,v)})\cdot d=\log\left((1+\frac{P\cdot\ell(d)}{N_0+I(u,v)})^d\right)$$

Clearly, when we fix the position of receiver v, to increase or decrease d by removing sender v will not change $I(u,v)$. When $0<d\le 1$, $l(d)=\min\{1,d^{-\beta}\}=1$, thus $\lambda^T(u,v)$ will get to its maximum value when $d=1$.

When $d>1$, $\lambda^T(u,v)=\log\left((1+\frac{P}{N_0+I(u,v)}d^{-\beta})^d\right)$. It is not difficult to show that $\log(1+\frac{P}{N_0+I(u,v)}d^{-\beta})^d$ is decreasing when $d>1$ where $\beta>2$. This finishes the proof. ■

Lemma 11: For an arbitrary network, when side length a of the deployment square is $a=\Omega(1)$ and $a=O(\sqrt{n})$, the total unicast throughput capacity for n transceiver pairs is $\Theta(a^2)$. In addition, the transport capacity is also $\Theta(a^2)$.

Proof: First, we prove that the total throughput capacity is at least $\Omega(a^2)$.

We partition the whole square into a^2 cells with side length 1. Next, we assume there is one transceiver pair in each cell. Assume for cell S_i, node u and v are chosen as source and receiver respectively. Based on a TDMA scheduling scheme, we let the transmitter in S_i be able to transmit only if all transmitters in the S_i and S_i's nearest 24 neighbor cells keep silent. Next we show that when all transceiver pairs in all grey cells exchange data simultaneously, for any pair of transmitter u and receiver v, the data rate between them is $\Omega(1)$ due to Lemma 5.

Next, we give a matching upper bound so that our results are indeed tight. First, we partition the whole square region into $\Theta(a^2)$ cells with side length $\Theta(1)$. Here, the partition is arbitrary, that means the partition can be shifted wherever you want. For those cells close to the boundary with side length $o(1)$, we simply consider them as a cell with side length $\Theta(1)$ as well, this will not affect the number of our cells ($\Theta(a^2)$), thus will not affect our proof. For any cell S_i, assume there are j simultaneously transmitting nodes $\{v_{i1},v_{i2},\cdots,v_{ij}\}$ inside S_j. Thus the unicast throughput capacity contributed by cell S_i is $\sum_{k=1}^{j}\lambda_{ik}$. Here, λ_{ik} is the feasible transmitting rate of the k^{th} transmitter inside of S_i. In addition, we do not assume that a receiver exists in the same cell as its transmitter. Clearly, adding one or more transmitters into S_i or replacing current transmitter(s) with others (originally silent nodes in this time slot) will not improve the unicast throughput capacity contributed by S_i due to Lemma 9. Thus, the total unicast throughput capacity is equal to $\sum_{i=1}^{\Theta(a^2)}\sum_{k=1}^{j}\lambda_{ij}$, which is bounded by $O(a^2\times 1)=O(a^2)$.

Observe that the total throughput capacity $\Omega(a^2)$ clearly is achievable by carefully placing a pair of nodes with distance $\Theta(1)$ in each cell. This construction also gives us a lower bound $\Omega(a^2)$ on the transport capacity. We then show that the transport capacity is also $O(a^2)$. Assume that $(u_1,v_1),(u_1,v_1),\cdots,(u_m,v_m)$ are m pairs of source-destinations that achieve the best transport capacity. Lemma 10 shows that the largest transport capacity will be achieved when each link has distance $O(1)$. This finishes the proof. ■

C. When $a=\Omega(\sqrt{n})$

Lemma 12: For an arbitrarily network, when the side length of square $a=\Omega(\sqrt{n})$, the total unicast throughput capacity for n transceiver pairs is $\Theta(n)$.

Proof: Clearly, the upper bound here is $O(n)$ because there are at most $\lfloor\frac{n}{2}\rfloor$ nodes that can transmit simultaneously with a constant transmission rate. Next we show that by the following construction, for an arbitrary wireless network, the total unicast throughput capacity is $\Omega(n)$ when the side length $a=\Omega(\sqrt{n})$.

We partition the region into $m=\frac{n}{c^2}$ small rectangles with side length $r=c\times\frac{a}{\sqrt{n}}$. Here, we can round c up to some constant such that $c^2\ge 2$ and m is an integer. In each small square, we put one source/destination pair within small distance d_1 around the center. First, we show when all transmitters transmit simultaneously, for any transceiver pair (u,v), the data rate is $R(u,v)=\Omega(1)$ based on a TDMA schedule. The proof idea is exactly same with the one used in Lemma 11. The total interference is

$$I(u,v)\le\sum_{i=1}^{\infty}8iP\cdot((2i-1)\frac{ca}{\sqrt{n}})^{-\beta}.$$

Notice that this sum clearly converges if $\beta>2$ when $a=\Omega(\sqrt{n})$, so $I(u,v)$ is a constant. Therefore, the data rate between u and v is $R(u,v)=\log(1+\frac{P\cdot\ell(d(u,v))}{N_0+I(u,v)})=\Omega(1)$, because $\ell(d(u,v))=\min\{1,|uv|^{-\beta}\}$ is also a constant.

Thus, at any time, based on our TDMA scheduling, there are at least $\lfloor\frac{n/2}{9}\rfloor$ links that can be active simultaneously, so the lower bound capacity for n transceiver pair is $\Omega(n)$. This finishes the proof. ■

In summary, from Lemmas 9, 11, and 12, we get Theorem 1.

IV. Unicast Capacity for Random Networks

In this section, we present both lower bounds and upper bounds of unicast capacity for a large scale random wireless network. We will study the capacity based on three scenarios $a=O(1)$, $a=\Omega(1)$ and $a=O(\sqrt{n})$, or $a=\Omega(\sqrt{n})$.

A. When $a = O(1)$

When the side length $a = O(1)$ and n goes to ∞, the unicast case for random wireless networks is similar with the one for arbitrary wireless networks.

Theorem 13: For a wireless network with n nodes randomly placed in a square region with side length a, the total unicast throughput capacity is $\Theta(1)$ when $a = O(1)$.

Proof: According to Lemma 9, we know the capacity for random wireless networks is also $O(1)$ when $a = O(1)$. In addition, if we only choose one pair of source/destination to transmit at any time slot the unicast capacity we can archive is $\Omega(1)$ by Lemma 9. This finishes the proof. ∎

Similarly it is not difficult to derive the following theorem.

Theorem 14: For a random wireless network with n randomly placed nodes in a square region B_a, the per-flow unicast throughput capacity is $\Theta(\frac{1}{n})$ when $a = O(1)$ and $n_s = \Theta(n)$.

B. When $\Theta(1) \leq a \leq \Theta(\sqrt{n})$

Next we show that when the side length a satisfies $1 \leq a \leq \sqrt{n}$, the per flow unicast throughput capacity is $\Omega(\frac{a}{n})$ when there are n unicast flows by constructing the following routing and link scheduling scheme.

By the percolation theory and the results in [4], when we partition the whole square into small cells with side length c, we can select one node from each cell and construct $\Omega(m)$ horizontal and $\Omega(m)$ vertical "highways" (or say disjoint paths) from left to right and from top to bottom respectively as the backbone of the whole wireless network with probability $1 - e^{-c^2}$. Here, $m = \frac{a}{c}$, where c is rounded up such that m is an integer. In addition, we can choose c large enough such that $\Omega(m)$ paths can be partitioned into a number of disjoint groups each with $\lceil \delta \log m \rceil$ disjoint paths, and each group are contained in a stripe with width m cells and height $(\kappa \log m - \epsilon_m)$ cells, for all $\kappa > 0$, δ small enough, and a non-zero small ϵ_m such that the side length of each stripe is integer. The same is true when we partition the square into vertical stripes with side length $m \times (\kappa \log m - \epsilon_m)$.

Routing Strategy: Our routing strategy is based on the backbone (highway system) we constructed. For each pair of source/destination nodes u and v, assume u is in the i^{th} stripe. If u is not on the highway, we will find a highway node u_{en} in the same stripe to be the entrance node of u, *i.e.*, u_{en} will be the first highway node which will relay packets of u. To find this entrance node, we draw a vertical line from u, and the closest highway node (from this line) which is in the same stripe will be chosen as u_{en}. For destination node v, if v is not a highway node, we use the same method to draw a vertical line from v, and find the closest highway node as the exit node u_{ex}. See Fig. 1 for illustration.

There are three phases for any pair of source/destination nodes (u, v) to communicate.

1) If u is not the highways nodes, u will find some entrance node u_{en} and send data to u_{en} by one hop.
2) u_{en} will relay the data of u to exit node v_{ex} of node v through highway (involving both vertical and horizontal highways).
3) v_{ex} will transmit the data to v directly at last.

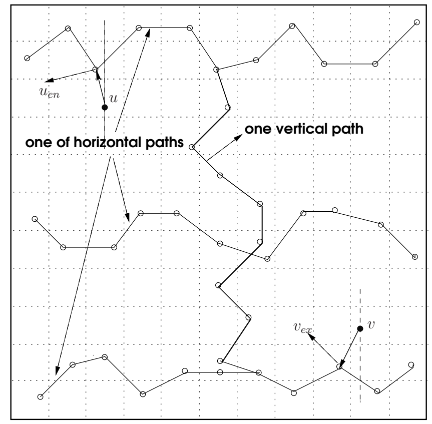


Fig. 1. For source/destination pair (u, v). u will draw a line and find entrance node u_{en} to highway and v will draw a line to find exit node v_{ex} from highway. This is also a simple case of our routing strategy for source/destination pair (u, v). u will send packet to u_{en} first, then $u_e n$ will relay packets to v_{ex} through both horizontal and vertical highway path. At last, v_{ex} will send packets to the destination node v.

Lemma 15: For any wireless node u, u can achieve a rate of $\Omega(\frac{a}{n})$ to some node v_{ex} on the highway system based on a TDMA schedule when the side length a satisfies $\Theta(1) \leq a \leq \Theta(\sqrt{n})$. Here $\log a > 1$.

Proof: We know that after we partition the whole square into horizontal (or vertical) stripes with size $m \times (\kappa \log m - \epsilon_m)$, node v can find an entrance node u_{en} on one of $\lceil \delta \log m \rceil$ disjoint paths within distance $\kappa \log m + \sqrt{2}c$ by the triangle inequality. By Lemma 5, we can get the data rate between u and u_{en} is $\Omega((\log m)^{-\beta-2})$. In addition, we know there are at most $\log m \times \frac{n}{m^2}$ nodes that will share the bandwidth together due to Lemma 6. Therefore the lower bound of the per-flow capacity is

$$\Omega(\frac{(\log m)^{-\beta-2}}{\log m \times \frac{n}{m^2}}) = \Omega(\frac{a}{n})$$

This finishes the proof. ∎

Clearly, the data rate achievable between destination node v and the exit node v_{ex} is $\Omega(\frac{a}{n})$ as well by applying Lemma 15 reversely. The remaining part is to compute the per-flow capacity of highway phase. Then we get our lower bound of unicast capacity by choosing the minimum one between draining phase and high-way phase.

Lemma 16: The nodes on the highways can achieve per-flow capacity rate of $\Omega(\frac{a}{n})$ with high probability based on a TDMA schedule when $\Theta(1) \leq a \leq \Theta(\sqrt{n})$.

Proof: By Lemma 5, because any two adjacent nodes on the highways are at most one cell away and the side-length c is a constant, any two adjacent nodes on highway can communicate with each other with constant rate based on a TDMA schedule.

In addition, we know that if we partition the square with side length a into $\frac{a}{c_1}$ stripes with size $c_1 \times a$, each stripe will contain at most $2\frac{c_1 n}{a}$ nodes *w.h.p.*, by Lemma 7. Here, c_1 can be rounded up such that $\frac{a}{c_1}$ is integer. Thus, for each node on the highway, it will relay traffic for at most $2\frac{c_1 n}{a}$ nodes with high probability. So, the per-flow capacity for each highway node is $\Omega(\frac{a}{n})$. This finishes the proof. ∎

Theorem 17: The per-flow unicast throughput capacity is $\Omega(\frac{a}{n})$ when $\Theta(1) \leq a \leq \Theta(\sqrt{n})$.

Proof: Based on Lemma 15 and Lemma 16, this theorem immediately follows when we choose the minimum one between the lower bound for draining phase and the lower bound for high way phase as the lower bound of per-flow unicast capacity. ∎

Next, by calculating a matching upper bound of per-flow unicast capacity, we can show that our results are indeed tight.

Lemma 18: Given a source/destination pair randomly placed in a square of side length a, the expected Euclidian distance between them is $c_2 a$ for some constant c_2.

Theorem 19: There is a constant c_3 such that, with probability at least $1 - 2e^{-n_s c_3^2/32}$, the per flow unicast throughput capacity that can be supported by any routing strategy is at most $\frac{c_3 a}{c n_s} = O(\frac{a}{n_s})$.
The detailed proof can be found in Appendix.

Obviously, when $a = \sqrt{n}$, our upper bound shows that the results in [4] is indeed tight.

C. When $a = \Omega(\sqrt{n})$

When we partition square region B_a into cells with side length g, as long as we scale g carefully, the highway system still exists. Notice that here g is not a constant but a function of a and n, $g = \theta_3 \frac{a}{\sqrt{n}}$, for some constant θ_3. In this case, we use the same routing strategy as described in Subsection IV-B to give a lower bound of unicast capacity first.

Lemma 20: For a random wireless network with n nodes randomly placed in B_a, by our routing strategy, the achievable per-flow unicast throughput capacity is $\Omega((\frac{a}{\sqrt{n}})^{-\beta} \cdot \frac{1}{\sqrt{n}})$ when $a = \Omega(\sqrt{n})$.

Proof: We partition the square B_a into $m^2 = (a/g)^2 = \theta_3^2 n$ cells with side length $g = \theta_3 \frac{a}{\sqrt{n}}$. Here we can choose constant θ_3 carefully such that $\frac{a^2}{g^2}$ is an integer. Then for any cell S_i, the probability that this cell S_i contains at least one node is $1 - e^{-\theta_3^2}$. Again, by appropriately choosing θ_3, we can make the above probability high enough. So the Euclidean distance between any two adjacent nodes (both horizontal and vertical) from the highway system we got by percolation theorem is bounded by $\sqrt{5}g$. Next, we use a TDMA scheduling such that a transmitter inside cell S_i can transmit iff all transmitters inside the closest 24 cells keep silent. Since any two adjacent highway nodes are at most one cell away from each other, then by Lemma 5, the transmission rate between any two adjacent nodes u and v on highway system is at least

$$\Omega(||g||^{-\beta}) = \Omega((\sqrt{5}c_5 \frac{a}{\sqrt{n}})^{-\beta}) = \Omega((\frac{a}{\sqrt{n}})^{-\beta}).$$

In addition, we know that each node on highway relays packets for at most $\frac{2c_1 n}{a} = O(\sqrt{n})$ nodes by Lemma 7. Thus, the per-flow unicast capacity is at least

$$\Omega((\frac{a}{\sqrt{n}})^{-\beta} \cdot \frac{1}{\sqrt{n}})$$

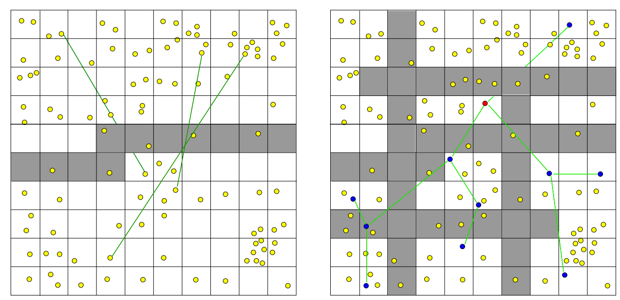
This finishes our proof. ∎

From now on, we will derive matching upper bound on the minimum per-flow unicast capacity. We first give the proof of the existence of a number of cells each of which contains only a constant number of nodes. Then we give an upper bound on minimum data rate by showing the congestion in those cells.

Definition 4: We say a cell is *quasi-closed cell* if it contains at most c_4 nodes, here c_4 is some constant. As illustrated in Figure. 2(a), we call a path of cells *quasi-closed cut* if it contains only quasi-closed cells and crosses from left to right side of B_a. Furthermore, we define the length of a quasi-closed cut as the total number of cells it contains.

Lemma 21: For any $\frac{5}{6} < p < 1$, there exists a constant c_4 such that the probability that any cell contains no more than c_4 nodes is at least p.

Proof: Denoting by x the number of nodes contained in one cell. We know that the expected value of x is $\frac{n}{a^2/(\theta_3 \cdot \frac{a}{\sqrt{n}})^2} = \theta_3^2$. According to Lemma 36, we get $\mathbf{Pr}(x \geq c_4) < \frac{c_4(1-\frac{\theta_3^2}{n})}{(c_4-\theta_3^2)^2}$. Thus, $\mathbf{Pr}(x < c_4) \geq 1 - \frac{c_4(1-\frac{\theta_3^2}{n})}{(c_4-\theta_3^2)^2}$. For any desired p, we can choose c_4 large enough such that $p \leq 1 - \frac{c_4}{(c_4-\theta_3^2)^2}$. This finishes our proof. ∎



(a) Quasi-closed cut (b) Quasi-closed cell net

Fig. 2. (a) Here a cell is called quasi-closed cell and marked grey if it contains at most c_4 nodes. (b) The colored cells compose a Quasi-closed cell net.

Lemma 22: Some quasi-closed cuts must be crossed by at least $c_5 n_s$ unicast sessions *w.h.p.*., for some constant c_5.

Proof: As shown in [4], for all $\kappa > 0$ and $\frac{5}{6} < p < 1$ satisfying $2 + \kappa \log(6(1-p)) < 0$, there exists a number of disjoint groups containing at least $\lceil \delta \log m \rceil$ disjoint paths in every group, and each group is constraint in a stripe of size $m \times (\kappa \log m - \epsilon_m)$ cells, for δ small enough satisfying

$$\delta \log \frac{p}{1-p} + 1 + \kappa \log(6(1-p)) < 1 \quad (7)$$

and a non-zero small ϵ_m such that the side length of each stripe is integer. Based on Lemma 21, same results can be used to prove the existence of our quasi-closed cuts.

For any constant $\kappa \leq \frac{1}{3} \frac{m}{\log m}$ and $\delta \geq \frac{1}{\log m}$, we pick c_4 carefully based on Lemma 21 to make sure that the preceding inequality (7) is satisfied. Then, *w.h.p.*, there exists at least three disjoint groups containing at least one quasi-closed cut in each group, and every group is bounded by a stripe with width a and height at most $\frac{a}{3}$. Here we only focus on the middle group, for each unicast session, the probability that it must cross the same quasi-close cut in the middle group is no less than $\frac{1}{3}$.

Denote by y the number of unicast sessions which cross the same quasi-closed cut belonging to the middle group.

According to Lemma 36, we get

$$\mathbf{Pr}\left(y \leq \frac{n_s}{6}\right) \leq e^{\frac{-2(\frac{n_s}{3}-\frac{n_s}{6})^2}{n_s}}$$

Thus, $\mathbf{Pr}\left(y > \frac{n_s}{6}\right) > 1 - e^{-\frac{n_s}{18}}$. Here if we set c_5 as $\frac{1}{6}$, the lemma follows when n_s goes to infinity. ■

Lemma 23: With high probability, some of the quasi-closed cells must be crossed by at least $c_6 \frac{n_s}{\sqrt{n}}$ unicast sessions for some constant c_6.

Proof: First, we will prove that, *w.h.p.*. in each group, there exists a quasi-closed cut whose length is no more than $\Theta(\sqrt{n})$. Since there are at least $\lceil \delta \log \frac{a}{g} \rceil$ disjoint paths in each group, and the size of one group is $\frac{a}{g} \times (\kappa \log \frac{a}{g} - \epsilon_{\frac{a}{g}})$, then by pigeonhole principle, there exists at least one quasi-closed cut, say Q, in each group which occupies no more than

$$\frac{\frac{a}{g} \times (\kappa \log \frac{a}{g} - \epsilon_{\frac{a}{g}})}{\lceil \delta \log \frac{a}{g} \rceil} = O(\frac{a}{g}) = O(\sqrt{n})$$

cells, when $g = \Theta(\frac{a}{\sqrt{n}})$. Then together with Lemma 22, there exists at least one cell in Q which is crossed by at least $c_6 \frac{n_s}{\sqrt{n}}$ unicast sessions for some constant c_6. Notice that it equals to $\Theta(\sqrt{n})$ when $n_s = \Theta(n)$. This finishes our proof. ■

Lemma 24: For a random wireless network with n nodes randomly placed in B_a, the minimum per-flow unicast capacity is at most $O((\frac{a}{\sqrt{n}})^{-\beta} \cdot \frac{1}{\sqrt{n}})$ when $a = \Omega(\sqrt{n})$.

Proof: According to Lemma 23, we know that for any routing strategy, there always exist some cells which contain only constant number of nodes while helping at least $c_6\sqrt{n}$ unicast sessions to relay (when $n_s = \Theta(n)$). Then the per-flow unicast capacity is at most

$$O(\frac{(\frac{a}{\sqrt{n}})^{-\beta}}{c_6\sqrt{n}}) = O((\frac{a}{\sqrt{n}})^{-\beta} \cdot \frac{1}{\sqrt{n}})$$

This finishes the proof. ■

Lemma 20 and Lemma 24 together imply Theorem 25.

Theorem 25: For a random wireless network with n nodes randomly placed in B_a with $a = \Omega(\sqrt{n})$, the minimum per flow unicast throughput capacity is $\Theta((\frac{a}{\sqrt{n}})^{-\beta} \cdot \frac{1}{\sqrt{n}})$.

Theorem 2 then follows from Theorem 14, Theorem 17, Theorem 19 and Theorem 25.

V. Multicast Capacity for Random Networks

In this section, we focus on deriving upper bounds for the multicast capacity. To study the multicast capacity, we first present one technique lemma which will be frequently used throughout this section.

Lemma 26: We partition square region B_a into cells with side length g. Given a multicast session M_i, let T_i be the multicast tree for M_i and $C(T_i)$ denote the number of cells used by the multicast tree T_i, then when $k < \theta_4 \cdot \frac{a^2}{g^2}$, *w.h.p.*, $C(T_i) \geq \theta_3 \frac{|EMST_{M_i}|}{g}$ where $|EMST_{M_i}|$ denotes the length of Euclidean Minimum Spanning Tree spanning M_i, θ_3 and θ_4 are some constants.

The detailed proof can be found in Appendix.

A. Upper Bound When $a = O(\sqrt{n})$

In this subsection, we provide an upper bound of multicast capacity when $a = O(\sqrt{n})$. Similar as previous approach, we partition the square region B_a with side length a into cells with side length c where c is some constant, then the total number of cells is $m^2 = \frac{a^2}{c^2} = \Theta(a^2)$.

Lemma 27: Given one multicast session M_i with one source and $k-1$ receivers randomly selected and all receivers are placed in a square region of side-length a, the Euclidean minimum spanning tree $\mathrm{EMST}(M_i)$ has an expected total edge length $c_1\sqrt{k}a$ for a constant $c_1 \in (0, 2\sqrt{2}]$

Theorem 28: When $a = O(\sqrt{n})$, with probability at least $1 - 2e^{-n_s c_8^2/32}$, the minimum per flow multicast throughout capacity by any routing strategy, is at most

$$O(\frac{a}{n_s\sqrt{k}}) \tag{8}$$

The detailed proof can be found in Appendix.

B. Upper Bound When $a = \Omega(\sqrt{n})$

Our main idea for upper bound on capacity is to show the existence of *quasi-closed cell net i.e.*, the cell net which is composed by all quasi-closed cells. Furthermore, by proving that with high probability, any multicast routing tree will cross a sufficient large number of quasi-closed cells, we can show that some cell will be used by many flows *i.e.*, the congestion in some quasi-closed cells. Please see Figure. 2(b) for illustration.

Next, we explain our proof in details: First we partition the square region B_a into $m^2 = c_8^2 n$ cells with side length $c_8 \frac{a}{\sqrt{n}}$ for some constant c_8. Then based on the results in [4] and Lemma 21, we can choose c_8 large enough such that $\Omega(m)$ quasi-closed cuts can be partitioned into a number of disjoint groups each with $\lceil \delta \log m \rceil$ disjoint quasi-closed cuts, and each group is constraint in a stripe of size $m \times (\kappa \log m - \epsilon_m)$, for all $\kappa > 0$, δ small enough, and a non-zero small ϵ_m such that the side length of each stripe is integer. The same is true when we partition the square into vertical stripes with side length $m \times (\kappa \log m - \epsilon_m)$. Notice that all of the horizontal and vertical stripes together partition B_a into *super-cells* with side length

$$(\kappa \log m - \epsilon_m) \times \frac{a}{m} = (\kappa \log m - \epsilon_m) \times c_8 \cdot \frac{a}{\sqrt{n}}$$

Theorem 29: When $k = O(\frac{n}{\lg^2 n})$ and $n_s = \Theta(n)$, with probability at least $1 - 2e^{-n_s c_8^2/32}$, the minimum per flow multicast throughout capacity by any routing strategy is at most $O((\frac{a}{\sqrt{n}})^{-\beta} \cdot \frac{1}{n_s} \cdot \frac{\sqrt{n}}{\sqrt{k}})$.

The detailed proof can be found in Appendix.

Observe that our result matches the upper bound derived in [5] when $a = \sqrt{n}$. Similar to the Theorem 2 of [5], we will derive another upper bound on multicast capacity using different approaches. The basic idea is to show that, for a random network topology, a cluster of nodes exists, that is relatively isolated from the rest of the nodes. The separation distance is at the same order of the size of the isolated cluster of nodes. Consequently, the average rate of the information that can be sent/received by the nodes of the cluster is limited.

Theorem 30: Assume that n_s random multicast flows are generated. The per-flow multicast throughout capacity $\varphi_k(n)$ is at most $O((\frac{a}{\sqrt{n}})^{-\beta}(\log n)^{1-\frac{\beta}{2}}/(n_s p_3))$, which is

$$\varphi_k(n) = \begin{cases} O(\frac{n}{n_s k}(\frac{a}{\sqrt{n}})^{-\beta}(\log n)^{-\frac{\beta}{2}}), & \text{if } k \le \frac{n}{\log n} \\ O(\frac{1}{n_s}(\frac{a}{\sqrt{n}})^{-\beta}(\log n)^{1-\frac{\beta}{2}}), & \text{if } k \ge \frac{n}{\log n} \end{cases} \quad (9)$$

Proof: Similar to the Theorem 2 of [5], we can prove that, with high probability, there is a cluster of $\Theta(\log n)$ nodes inside a cell of size $h = \frac{a}{3}\sqrt{\frac{\log n}{n}}$, and the cluster is separated from the rest of nodes with distance at least $\frac{a}{3}\sqrt{\frac{\log n}{n}}$. Furthermore let p_3 be the probability that at least one node in this isolated cluster is a source or terminal of a multicast session. Then we can show that $p_3 = \Theta(\frac{k \log n}{n})$ when $k = O(\frac{n}{\log n})$, and $p_3 = O(1)$ when $k = \Omega(\frac{n}{\log n})$. Obviously, the maximum link rate that can be supported for any link uv with u inside this isolated cluster and v outside of this cluster is at most $\log(1 + \frac{P \cdot h^{-\beta}}{N_0}) = \Theta(h^{-\beta})$ since $h \to \infty$ when $n \to \infty$. Notice that there are $\Theta(\log n)$ nodes inside this isolated cluster. Consequently, the total data rate that can be transmitted from/to this cluster is at most $\Theta(\log n) \cdot \Theta(h^{-\beta}) = \Theta((\frac{a}{\sqrt{n}})^{-\beta}(\log n)^{1-\frac{\beta}{2}})$ since each node inside the cluster can not receive from multiple nodes. Since there are n_s flows and the probability that a given flow has a receiver node inside this isolated cluster is p_3, the expected number of flows that will have receivers inside this isolated cluster is $p_3 n_s$. Using Azuma's inequality we are able to prove that with high probability, there are $p_3 n_s/2$ flows that will have receivers inside this isolated cluster. Let φ be the minimum per-flow multicast data rate. Thus, we have

$$\varphi \cdot p_3 n_s/2 \le c_{11}(\frac{a}{\sqrt{n}})^{-\beta}(\log n)^{1-\frac{\beta}{2}}$$

for some constant c_{11}. Consequently

$$\varphi = O((\frac{a}{\sqrt{n}})^{-\beta}(\log n)^{1-\frac{\beta}{2}}/(n_s p_3)).$$

The theorem then follows directly. ■

The preceding upper bound on multicast is derived by analyzing an isolated cluster of nodes. For a random wireless network (n nodes randomly distributed in a region B_a, or nodes following a Poisson distribution with parameter $\frac{n}{a^2}$), it is proved in [13] that, *w.h.p.*, the nearest neighbor graph has an edge of length $\Theta(a\sqrt{\frac{\log n}{n}})$. By exploring this long edge, we are able to derive another upper bound on multicast capacity.

Theorem 31: The per-flow multicast throughout capacity $\varphi_k(n)$ of n_s flows, when deployment square B_a has side-length $a = \Omega(\sqrt{n})$ and $k = \omega(\sqrt{n})$, is at most

$$\varphi_k(n) = O\left(\frac{1}{n_s}\frac{n}{k}\left(\frac{a}{\sqrt{n}}\right)^{-\beta}(\log n)^{-\frac{\beta}{2}}\right) \quad (10)$$

Proof: Assume that the longest edge in the nearest neighbor graph of the random network is uv. Then for node v, the probability p_3 that it is chosen as a terminal of a given multicast flow is $p_3 = \frac{k}{n}$. It is easy to show that, with high probability (at least $1 - e^{-\frac{k^2}{2n}}$), the number of multicast flows that will choose the node v as a terminal is at least $n_s p_3/2$ when $k = \omega(\sqrt{n})$. Observe that the total data rate that node v can receive is at most $R(v) = O\left(\left(\frac{a}{\sqrt{n}}\right)^{-\beta}(\log n)^{-\frac{\beta}{2}}\right)$ since the shortest link incident at node v is at least $\Theta(a\sqrt{\frac{\log n}{n}})$. Then we have $\varphi_k(n) \cdot n_s p_3/2 \le R(v)$. The theorem then directly follows. ■

Combining Theorem 29, Theorem 30, and Theorem 31, we have Theorem 4. When $n_s = n$ and $a = \sqrt{n}$, we have

Corollary 32: The per-flow multicast throughout capacity $\varphi_k(n)$ of n flows for networks in $B_{\sqrt{n}}$ is at most

$$\varphi_k(n) = \begin{cases} O(\frac{1}{\sqrt{n}\sqrt{k}}) & \text{if } k \le \frac{n}{(\log n)^{\beta}} \\ O(\frac{1}{k}(\log n)^{-\frac{\beta}{2}}), & \text{if } k \ge \frac{n}{(\log n)^{\beta}} \end{cases} \quad (11)$$

In a summary, we can get Theorem 3 based on above discussions. Observe that our upper bound on multicast capacity is achievable when $k = n$ for broadcast [19] and $k = O(\frac{n}{(\log n)^{2\beta+6}})$ [20]. Our upper bounds also improve the result in [5].

VI. Literature Reviews

Gupta and Kumar [3] studied the asymptotic 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 fixed 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 [3] 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 [7] recently showed that mobility actually can help to improve the 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 [7] 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 [7] still obey the capacity bound proposed in [3]: the capacity is improved because the average distance $\overline{L}$ a packet has to be transmitted is reduced from $\Theta(1)$ in [3] to $\Theta(r(n))$ in [7]. In summary, for random networks, under the protocol model, the achievable throughput capacity $\lambda(n)$ and the average travel distance $\overline{L}$ satisfies

$\lambda(n) \cdot \overline{L} \leq \Theta(\frac{W}{\Delta^2 n \cdot r(n)})$. This phenomenon has also been observed in [10]. They found that the traffic pattern determines whether the per node 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 node 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.

In [3], the capacity of wireless networks are solved under a number of assumptions, among them point-to-point coding which excludes for example the multi-access and broadcast codes. In [6] Gastpar and Vetterli studied the capacity of wireless networks when network coding can be used to improve the capacity. They essentially considered the same physical model under different traffic pattern (relay traffic pattern). They allow for arbitrary complex network coding. In their model, there is only one source and destination pair while all other nodes will assist this transmission. They show that the capacity of such wireless networks with n nodes under relay traffic pattern behaves like $\log n$ bits per second. This demonstrates the power of network coding: under the point-to-point coding assumption considered in [3], the achievable data rate is constant, independent of the number of nodes.

Broadcast capacity of an arbitrary network has been studied in [9], [14]. They essentially show that the broadcast capacity of a given network is $\Theta(W)$ for single source broadcast and the achievable broadcast capacity per node is only $\Theta(W/n)$ if each of the n nodes will serve as source node.

Multicast capacity was recently studied in the literature, *e.g.*, [11], [12], [15], [16], just name a few. Jacquet and Rodolakis [16] studied the scaling properties of multicast for random wireless networks. They claimed that the maximum rate at which a node can transmit multicast data is $O(\frac{W}{kn \log n})$. Li *et al.* [11], [12] studied asymptotic multicast capacity for protocol interference model. Assume for each node v_i $(1 \leq i \leq n)$, randomly and independently pick $k-1$ points $p_{i,j}$ $(1 \leq j \leq k-1)$ from the square, and then v_i multicast data to the nearest node for each $p_{i,j}$. They defined the aggregated multicast capacity as the total data rate of all multicast sessions in the network and then gave the matching asymptotic upper bounds and lower bounds for it. They showed the total multicast capacity is $\Theta(\sqrt{\frac{n}{\log n}} \cdot \frac{W}{\sqrt{k}})$ when $k = O(\frac{n}{\log n})$ and when $k = \Omega(\frac{n}{\log n})$, the total multicast capacity is equal to the broadcast capacity, *i.e.*, $\Theta(W)$. Mao et al. [18] studied the multicast capacity for hybrid networks. They derived several capacity regimes based on the relations of the number k of receivers per multicast session, the total number n of nodes, and the number m of base stations.

Instead of studying the capacity problem either under protocol model or physical model, Franceschetti et al. [4] addressed the unicast capacity under Gaussian Channel, they proposed a routing and scheduling scheme using high way system based on percolation theorem and proved that a rate $\frac{1}{\sqrt{n}}$ is achievable in networks of randomly located nodes when Gaussian channel is used. Zheng [19] pointed out that using multihop relay, the rate of broadcasting continuous stream is $\Theta((\log n)^{-\frac{\beta}{2}})$ in *random extended networks*. Regardless of the density, information can diffuse at constant speed, i.e., $\Theta(1)$. Most recently, Li et al. [20] proposed that, when $n_d = O(\frac{n}{(\log n)^{2\beta+6}})$ and $n_s = \Omega(n^{\frac{1}{2}+\theta})$, the achieving per-session multicast throughput is w.h.p. of order $\Omega(\frac{\sqrt{n}}{n_s \sqrt{n_d}})$ using *percolation model*, where $\theta > 0$ is a constant. All the above results are derived under the *bounded propagation model* ([21]) and for a single network. Gupta et al [22] study the transport capacity of the Gaussian multiple access channel (MAC), which consists of multiple transmitters and a single receiver, and the Gaussian broadcast channel (BC), which consists of a single transmitter and multiple receivers. Recently, Toumpis [17] study the capacity bounds when the traffic is non-uniform, i.e., not all nodes are required to receive and send similar volumes of traffic. Their results are asymptotic, i.e., they hold with probability going to unity as the number of nodes goes to infinity. Keshavarz-Haddad and Riedi [5] derived some upper bounds on multicast capacity for Gaussian channel model. They also present algorithms for multicast and claimed that the capacity achieved by their method matches the upper bound. Their bounds are not tight, *e.g.*, rate of W is not achievable when $k \geq \frac{n}{\log n}$.

VII. Conclusions

In this paper, we studied the impact of the size of the deployment region on the asymptotic unicast and multicast throughput capacity for wireless networks under Gaussian channel model. We derived asymptotic matching upper-bounds and lower-bounds of unicast capacity for arbitrary and random wireless networks in different cases. Our upper-bounds improve the previous result and uses new analyzing techniques. A number of interesting and challenging questions remain as future work. The first question is to close the gap on multicast capacity by presenting possibly new tight upper-bounds and designing algorithms to achieve the asymptotic multicast capacity. Next, what will be the multicast capacity when nodes have different transmission power, or each node can adjust the transmission power during the communication?

References

[1] S. Kulkarni and P. Viswanath A deterministic approach to throughput scaling in wireless networks. IEEE Trans. Inf. Theory, vol. 50, pp. 1041-1049, 2004

[2] S. Toupis and A.J. Goldsmith Large wireless networks under fading, mobility, and delay constraints. Proc. IEEE Trans. Inf. Commun. Conf., INFOCOM 2004.

[3] Gupta, P., and Kumar, P. Capacity of wireless networks. Tech. rep., University of Illinois, Urbana-Champaign, 1999.

[4] Massimo Franceschetti, Olivier Dousse, David N.C. Tse, and Patrick Thiran Closing the Gap in the Capacity of Wireless Networks Via Percolation Theory. In *IEEE Transactions on Information Theory Vol. 53.* (2007).

[5] Alireza Keshavarz-Haddad, Rudolf Riedi. Multicast Capacity of Large Homogeneous Multihop Wireless Networks, 6th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2008.

[6] Gastpar, M., and Vetterli, M. On the capacity of wireless networks: the relay case. In *IEEE INFOCOM* (2002).
[7] Grossglauser, M., and Tse, D. Mobility increases the capacity of ad-hoc wireless networks. In *INFOCOMM* (2001), vol. 3, pp. 1360 –1369.
[8] Gupta, P., and Kumar, P. R. 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).
[9] Keshavarz-Haddad, A., Ribeiro, V., and Riedi, R. Broadcast capacity in multihop wireless networks. In *ACM MobiCom* (2006), pp. 239–250.
[10] Li, J., Blake, C., Couto, D. S. J. D., Lee, H. I., and Morris, R. Capacity of ad hoc wireless networks. In *ACM MobiCom* (2001).
[11] Xiang-Yang Li Multicast Capacity for Large Scale Wireless Ad Hoc Networks *IEEE Trans. on Networking* (May. 2008).
[12] Li, X.-Y., Tang, S.-J., and Frieder, O., Multicast Capacity of Large Scale Wireless Ad Hoc Networks, In *ACM MobiCom* (2007).
[13] Penrose, M. The longest edge of the random minimal spanning tree. *Annals of Applied Probability 7* (1997), 340–361.
[14] Tavli, B. Broadcast capacity of wireless networks. *IEEE Communication Letters 10*, 2 (February 2006).
[15] Shakkottai S., Liu X. and Srikant R., The multicast capacity of ad hoc networks, *Proc. ACM Mobihoc*, 2007.
[16] Jacquet, P. and Rodolakis, G., Multicast scaling properties in massively dense ad hoc networks, *ICPADS*, 2005.
[17] Stavros Toumpis, Asymptotic Capacity Bounds for Wireless Networks with Non-Uniform Traffic. IEEE transactions on wireless communications, 2008, vol. 7, no6, pp. 2231-2242.
[18] Mao, X., Li, X.-Y., AND Tang, S.-J., Multicast capacity for hybrid wireless networks. *ACM MobiHoc (2008).*
[19] R. Zheng, Asymptotic bounds of information dissemination in power-constrained wireless networks, *IEEE Trans. on Wireless Communications*, vol. 7, no. 1, pp. 251–259, Jan. 2008.
[20] S. Li, Y. Liu, and X.-Y. Li, Capacity of large scale wireless networks under gaussian channel model, in *ACM Mobicom*, 2008.
[21] A. Keshavarz-Haddad and R. Riedi, Bounds for the capacity of wireless multihop networks imposed by topology and demand, in *ACM MobiHoc'07*, 2007.
[22] Gupta, G.A. and Toumpis, S. and Sayir, J. and Müller, R.R. On the transport capacity of Gaussian multiple access and broadcast channels in *WINET*, 2006.

VIII. Appendix

A. Useful Known Results

Throughout this paper, we will repeatedly use the following results from probability theory literature.

Lemma 33 (Azuma's Inequality): Suppose that random variables X_0, X_1, X_2, $\cdots$, X_n, $\cdots$ are martingale and $|X_k - X_{k-1}| \le a_k$ almost surely for any $k \ge 1$. Then for all positive integers N and all positive real number t, we have

$$\mathbf{Pr}\left(|X_N - X_0| \ge t\right) \le 2\exp\left(-\frac{t^2}{2\sum_{i=1}^{N} a_k^2}\right)$$

Recall that here a sequence of random variables X_i, $0 \le i$, are called *martingale* if they satisfy that: $E(X_{N+1} \mid X_0, X_1, \cdots, X_N) = X_N$.

Lemma 34 (Chebyshev's Inequality): For a variable X,

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

where $\mu = E(X)$, $Var(X)$ is the variance of X, and $A > 0$.

Lemma 35 (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 $\sigma^2 = Var(X_i)$. Let $\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$. $\forall \epsilon > 0$,

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

Lemma 36 (Binomial Distribution): Consider n independent variables $X_i \in \{0,1\}$, $p = \mathbf{Pr}\left(X_i = 1\right)$, and $X = \sum_{i=1}^{n} X_i$.

$$\begin{cases} \mathbf{Pr}\left(X \le \xi\right) \le e^{\frac{-2(n\cdot p-\xi)^2}{n}}, & \text{when } 0 < \xi \le n\cdot p. \\ \mathbf{Pr}\left(X > \xi\right) < \frac{\xi(1-p)}{(\xi-n\cdot p)^2}, & \text{when } \xi > n\cdot p. \end{cases} \quad (12)$$

Lemma 37: [4] For a Poisson random variable X of parameter λ,

$$\mathbf{Pr}\left(X \ge x\right) \le \frac{e^{-\lambda}(e\lambda)^x}{x^x}, \text{ for } x > \lambda$$

B. Proof of some technical lemmas

Proof for Lemma 5:

Proof: We use the similar idea as Theorem 3 in [4] to prove this. We first partition the square into $\frac{a^2}{c^2}$ cells with side length c, here c can be rounded such that $\frac{a^2}{c^2}$ is an integer. Then we divide time into a sequence of $q = (c_3(d+1))^2$ successive mini-time-slots. Here, c_3 is a constant no less than 2 and can be rounded such that q is an integer. Then based on a TDMA schedule, in each mini-time-slot, we let only one node in each of the disjoint sets (square with color grey) to transmit simultaneously.

First, we know the distance between u and v is at most $\sqrt{2}(d+1)c$. Hence, the signal strength received by v is at least $S(u,v) \ge P\cdot \ell(\sqrt{2}(d+1)c) = P\cdot\min\{1, (\sqrt{2}(d+1)c)^{-\beta}\}$. Next,we analyze the total interference received by v based on our TDMA scheduling described above. Given a transmitter u in one cell s_i, we know the receiver v is in some cell that is at most d-cells apart from s_i. The total interference caused by all other simultaneous transmissions can be computed as follows. First, any transmitter located in one of the eight closest cells is at least $c_3(d+1) - (d+1)$ cells away from v, then the Euclidean distance is at least $c\cdot(c_3(d+1)-(d+1))$. Next, any transmitter of the sixteen next closest cells has Euclidean distance from v is at least $c\cdot(2c_3(d+1)-(d+1))$, and so forth. Thus, the total interference caused by all simultaneous transmitters is

$$\begin{aligned} I(u,v) &\le \sum_{i=1}^{\infty} 8i\cdot P\cdot(c\cdot i\cdot(c_3-1)d + i\cdot c_3 - 1)^{-\beta} \\ &\le \sum_{i=1}^{\infty}\frac{8P}{c^\beta}\frac{i}{(i(c_3-1)d)^\beta} = \frac{8P}{(c_3-1)^\beta(cd)^\beta}\sum_{i=1}^{\infty}\frac{1}{i^{\beta-1}} \end{aligned}$$

Clearly, when $\beta > 2$, the summation in the above formula converges. So the total interference is at most $c_2 P\cdot(cd)^{-\beta}$, here $c_2 = \frac{8}{(c_3-1)^\beta}\sum_{i=1}^{\infty}\frac{1}{i^{\beta-1}}$ is a constant.

We can get the data rate $R(u,v)$ is at least

$$\begin{aligned} R(u,v) &= \log(1 + \frac{S(u,v)}{N_0 + I(u,v)}) \\ &\ge \log(1 + \frac{P\cdot\min\{1,(\sqrt{2}(d+1)c)^{-\beta}\}}{N_0 + c_2 P(cd)^{-\beta}}) \end{aligned}$$

which does not depend on n.

Clearly, when both c and d are constants, $R(u,v) = \Theta(1)$. When $c\cdot d \to \infty$, by taking the limit for $c\cdot d\to\infty$ and by the fact that every transmitter can transmit once every $q^2 = (c_3(d+1))^2$ mini-time-slots, the lemma follows. ■

Proof for Lemma 6:

Proof: The proof follows from Lemma 37. Let A_n be the event that there is at least one cell with more than $\log \frac{a}{c} \times \frac{nc^2}{a^2}$ nodes. Since the number of nodes x in each cell of the partition is a Poisson random variable of parameter $\frac{nc^2}{g^2}$, by the union the Chernoff bounds, we have $\Pr(A_n) \leq (\frac{a}{c})^2 \Pr(x > \log \frac{a}{c} \times \frac{nc^2}{a^2}) \leq (\frac{a}{c})^2 e^{-\frac{nc^2}{a^2}} (\frac{\frac{nc^2}{a^2} e}{\log \frac{a}{c} \times \frac{nc^2}{a^2}})^{\frac{nc^2}{a^2} \log \frac{a}{c}}$. which tends to 0 as n tends to infinity. ■

Proof for Lemma 7:

Proof: Let x be the number of nodes falling in one rectangle with size $c_1 \times a$ and A_n be the event that there is at least one rectangle with more than nodes, by Lemma 37, we get $\Pr(A_n) \leq \frac{\sqrt{n}}{c_1} \times \Pr(x > c_1\sqrt{n}) \leq \frac{\sqrt{n}}{c_1} e^{-c_1\sqrt{n}} (\frac{ec_1\sqrt{n}}{c_1\sqrt{n}})^{2c_1\sqrt{n}} = \frac{\sqrt{n}}{c_1} e^{-c_1\sqrt{n}} (\frac{e}{2})^{2c_1\sqrt{n}}$. When n tends to infinity, it goes to 0. This finishes the proof. ■

Proof for Theorem 19:

Proof: First, we partition B_a into cells with side length c, here c is some constant. Let $C(P_i)$ denote the number of cells a routing path P_i will use, *i.e.*, the number of cells crossed by P_i. Let variable $L = \sum_{i=1}^{n_s} C(P_i)$, denoting the total load of all cells. Here the load of a cell by a routing method is the number of flows visiting the cell for the unicast path constructed. Then $L \geq \sum_{i=1}^{n_s} l_i/(\sqrt{2}\frac{a}{m})$, where l_i denotes the Euclidian distance between the i-th source/destination pair.

Define random variables $X_q = \sum_{j=1}^{q}(l_j - E(l_j))$. Then $E(X_{q+1} \mid X_1, \cdots, X_q) = E(\sum_{j=1}^{q+1}(l_j - E(l_j)) \mid X_1, \cdots, X_q) = E(\sum_{j=1}^{q}(l_j - E(l_j)) + (l_{q+1} - E(l_{q+1}) \mid X_1, \cdots, X_q) = X_q + E(l_{q+1} - E(l_{q+1})) = X_q$. In other words, variables X_i are martingale.

In addition, $|X_q - X_{q-1}| = |l_q - E(l_q)| \leq \sqrt{2}a$. Note that the last inequality holds for any l_q [12]. From Azuma's Inequality, we have $\mathbf{Pr}(|X_{n_s} - X_0| \geq t) \leq 2\exp(-\frac{t^2}{2\sum_{i=1}^{n_s} 8a^2})$. Let $t = \epsilon \sum_{i=1}^{n_s} E(|l_i|)$. Clearly, $\epsilon n_s c_3 a \leq t \leq \epsilon n_s \sqrt{2a}$ for some constant c_3. Note that $X_0 = 0$. Then, $\mathbf{Pr}(\sum_{i=1}^{n_s} l_i \leq \sum_{i=1}^{n_s} E(l_i) - t) \leq \mathbf{Pr}(|X_{n_s}| \geq t) \leq \exp(-\frac{t^2}{2\sum_{i=1}^{n_s} 8a^2}) \leq \exp(-\frac{(\epsilon n_s c_3 a)^2}{8n_s a^2}) = \exp(-\frac{n_s \epsilon^2 c_3^2}{8})$. Thus, for a constant $\epsilon \in (0, 1)$,

$$\mathbf{Pr}\left(\sum_{i=1}^{n_s} l_i \leq (1-\epsilon) n_s \sqrt{2}a\right) \leq 2e^{-\frac{n_s \epsilon^2 c_3^2}{8}}$$

Thus, by letting $\epsilon = \frac{1}{2}$, we have $\mathbf{Pr}\left(\sum_{i=1}^{n_s} l_i \geq n_s\sqrt{2}a/2\right) \geq 1 - 2e^{-n_s c_3^2/32}$. Then, $\mathbf{Pr}(L \geq n_s m/2) \geq 1 - 2e^{-n_s c_3^2/32}$.

Recall that L denotes the total load of all cells. Then by pigeonhole principle, with probability at least $1 - 2e^{-n_s c_3^2/32}$, there is at least one cell, that will be used by at least $\frac{n_s c_3 m}{m^2}$ flows. According to Lemma 9, we know that the capacity of a cell with constant side length is $O(1)$. Thus, with probability at least $1 - 2e^{-n_s c_3^2/32}$, the minimum data rate that can be supported using any routing strategy, due to the congestion in some cell, is $\frac{1}{\frac{n_s m}{c_3 m^2}} = \frac{c_3 m}{n_s} = \frac{c_3 a}{c n_s} = O(\frac{a}{n_s})$, since $m = a/c$ for some constant c. ■

Proof for Lemma 26:

Proof: We will prove this lemma using some existing results under protocol model, especially the area argument [12]. For the sake of our proof, assume that every node has an artificial "transmission radius" r such that each node v can only communicate with other nodes in its transmission range (a transmitting disk with its center at v and radius r). In addition, we define the area covered by a tree T as the union of its nodes' transmitting disks. Then by showing a lower bound on the area of the region covered by any multicast tree T, we can give the desired lower bound on the number of cells it will cross.

Recall Lemma 11 in [11], it is proved that in protocol model, the area of the region $D(T)$, *w.h.p.*, is at least $\theta_0\sqrt{k}ar$ when $k < \theta_1 \cdot \frac{a^2}{r^2}$ for some constant θ_0 and θ_1. Here r denotes the transmission range of each node in protocol model and $D(T)$ denotes the region covered by all transmitting disks of all transmitting nodes (internal nodes of T) in the any multicast tree T. Unfortunately, this result can not help us directly, since in our model, each node has no fixed transmission range r. Instead, any pair of nodes can communicate with each other even though the data rate may be very small. Based on the original network under Gaussian channel model, we construct a new network under protocol model as follows.

1) Set each node's transmission range as the side length of each cell g.
2) Add some artificial "additional relay nodes" V_{ad} such that any pair of nodes will have enough relay nodes along its link to make sure that the minimum number of cells the routing path crosses under protocol model is no more than the number of cells the direct link will cross in Gaussian channel model. Notice that V_{ad} cannot be selected as source or receivers, they can only act as relay nodes.

Let T be any multicast tree in original network under Gaussian channel model and T_p denote the corresponding multicast tree(spanning the same multicast session) constructed on this network under protocol model. We have two important observations here:

1) Our preceding two modifications will not affect the proof for Lemma 11 in [11]. In other words, the lower bound on $|D(T_p)|$ still holds,
2) Furthermore, any link in Gaussian channel model can be simulated by using these artificial "additional relay nodes" in the protocol model such that the number of cells it will cross is not increased. So the lower bound of $C(T)$ is no smaller than the lower bound of $C(T_p)$.

Together with Lemma 11 in [11], we get

$$D(T_p) \geq \theta_0\sqrt{k}ar = \theta_0\sqrt{k}ag \tag{13}$$

Since one transmitting disk can cover no more than 4 cells, we have $C(T_p) \geq \theta_0\sqrt{k}ag/4 \times g^2 = \frac{\theta_0}{4} \cdot \frac{\sqrt{k}a}{g}$. It follows that when $k < \theta_4 \cdot \frac{a^2}{g^2}$, with high probability,

$$C(T) \geq \frac{\theta_0}{4} \cdot \frac{\sqrt{k}a}{g}.$$

Since $|EMST| \leq 2\sqrt{2}\sqrt{k}a$, if we set θ_3 as $\frac{\theta_0}{4}/2\sqrt{2}$, our lemma follows. ■

Proof for Theorem 28:

Proof: Let variable $L = \sum_{i=1}^{n_s} C(T_i)$, denoting the total load of all cells. Here the load of a cell by a routing method is the number of flows visiting the cell for the multicast tree constructed. Then based on Lemma 26, we know that $L \geq \sum_{i=1}^{n_s} \theta_3 |\operatorname{EMST}(M_i)|/(\frac{a}{m})$ with high probability. Notice that $E(\sum_{i=1}^{n_s} |\operatorname{EMST}(M_i)|) = n_s c_7 a\sqrt{k}$.

Define random variables $X_q = \sum_{j=1}^{q}(|\operatorname{EMST}(M_j)| - E(|\operatorname{EMST}(M_j)|))$. Then $E(X_{q+1} \mid X_1, \cdots, X_q) = E(\sum_{j=1}^{q+1}(l_j - E(l_j)) \mid X_1, \cdots, X_q) = E(\sum_{j=1}^{q}(l_j - E(l_j)) + (l_{q+1} - E(l_{q+1}) \mid X_1, \cdots, X_q) = X_q + E(l_{q+1} - E(l_{q+1})) = X_q$, so variables X_i are martingale.

$|X_q - X_{q-1}| = ||\operatorname{EMST}(M_q)| - E(|\operatorname{EMST}(M_q)|)| \leq E(|\operatorname{EMST}(M_q))| \leq 2\sqrt{2}\sqrt{k}a$. This inequality holds for any $\operatorname{EMST}(M_q)$ [12].

From Azuma's Inequality, we have $\mathbf{Pr}\left(|X_{n_s} - X_0| \geq t\right) \leq 2\exp(-\frac{t^2}{2\sum_{i=1}^{n_s} 8ka^2})$ Let $t = \epsilon\sum_{i=1}^{n_s} E(\operatorname{EMST}(M_i))$. Clearly, $\epsilon n_s c_8 \sqrt{k}a \leq t \leq 2\sqrt{2} n_s \epsilon \sqrt{k}a$. Note that $X_0 = 0$. Then, $\mathbf{Pr}\left(\sum_{i=1}^{n_s} |\operatorname{EMST}(M_i)| \leq \sum_{i=1}^{n_s} E(|\operatorname{EMST}(M_i)|) - t\right) \leq \mathbf{Pr}\left(|X_{n_s}| \geq t\right) \leq \exp(-\frac{t^2}{2\sum_{i=1}^{n_s} 8ka^2}) \leq \exp(-\frac{(\epsilon n_s c_8 \sqrt{k}a)^2}{8 n_s k a^2}) = \exp(-\frac{n_s \epsilon^2 c_8^2}{8})$. Thus, for a constant $\epsilon \in (0, 1)$,

$$\mathbf{Pr}\left(\sum_{i=1}^{n_s} |\operatorname{EMST}(M_i)| \leq (1-\epsilon) n_s c_9 \sqrt{k}a\right) \leq 2e^{-\frac{n_s \epsilon^2 c_8^2}{8}}$$

Then by letting $\epsilon = \frac{1}{2}$, we have

$$\mathbf{Pr}\left(\sum_{i=1}^{n_s} |\operatorname{EMST}(M_i)| \geq n_s c_9 \sqrt{k}a/2\right) \geq 1 - 2e^{-n_s c_8^2/32}.$$

Based on Lemma 26, we get

$$\mathbf{Pr}\left(L \geq n_s \theta_3 c_9 \sqrt{k} m/2\right) \geq 1 - 2e^{-n_s c_8^2/32}.$$

It implies that

$$\mathbf{Pr}\left(L \geq n_s \theta_3 c_9 \sqrt{k} m/2\right) \geq 1 - 2e^{-n_s c_8^2/32} \text{ if } k \leq \theta_1 \sqrt{n}. \quad (14)$$

Recall that L denotes the total load of all cells. Then by pigeonhole principle, with probability at least $1 - 2e^{-n_s c_8^2/32}$, there is at least one cell, that will be used by at least $\frac{n_s c_{10}\sqrt{k}m}{m^2}$ flows where $c_{10} = \theta_3 c_9$. Again, according to Lemma 9, the capacity of a cell with constant side length is $O(1)$. Thus, when $n_s = \Theta(n)$, with probability at least $1 - 2e^{-n_s c_8^2/32}$, the minimum data rate that can be supported using cellular routing strategy is at most, for any routing strategy, due to the congestion in some cell, $\frac{1}{\frac{n_s c_{10}\sqrt{k}m}{m^2}} = \frac{m}{c_{10} n_s \sqrt{k}} = O(\frac{a}{n_s\sqrt{k}})$. This finishes the proof of the theorem. ■

Proof for Theorem 29:

Proof: Our proof again is to analyze the load of some cells. We use L to denote the total load of all cells. Then we get $L \geq \sum_{i=1}^{n_s} \theta_3 |EMST(M_i)| / \left((\kappa \log m - \epsilon_m) \times \frac{a}{m}\right)$ based on Lemma 26. Since $\mathbf{Pr}\left(\sum_{i=1}^{n_s} |\operatorname{EMST}(M_i)| \geq n_s c_9 \sqrt{k}a/2\right) \geq 1 - 2e^{-n_s c_8^2/32}$, from Lemma 26, we get

$$\mathbf{Pr}\left(L \geq n_s c_{10} \sqrt{k} \frac{m}{\lg m}\right) \geq 1 - 2e^{-n_s c_8^2/32}.$$

for some constant $c_{10} = c_9 \theta_3$. Here we use $\mathbb{L}$ to denote the total number of flows crossing some super-cell. Notice that here "crossing" means visiting and leaving. We get,

$$\mathbf{Pr}\left(\mathbb{L} \geq L - n_s k = n_s c_{10} \sqrt{k} \frac{m}{\lg m}/2 - n_s k\right) \geq 1 - 2e^{-n_s c_8^2/32}.$$

It is easy to show that, any multicast routing tree will cross at least $\lceil \delta \lg m \rceil$ quasi-closed cuts if it crosses three super-cells. Denoted by $\mathbb{L}'$ the total number of flows crossing some quasi-closed cut. We have $\mathbb{L}' \geq \frac{\mathbb{L}}{3} \times \lceil \delta \lg m \rceil$.

It follows that, with probability at least $1 - 2e^{-n_s c_8^2/32}$, the total load of all quasi-closed cell is at least

$$\frac{n_s c_{10} \sqrt{k} \frac{m}{\lg m}/2 - n_s k}{3} \times \lceil \delta \lg m \rceil.$$

Then by pigeonhole principle, with probability at least $1 - 2e^{-n_s c_8^2/32}$, there is at least one quasi-closed cell, that will be used by at least $\frac{\frac{n_s c_{10} \sqrt{k} \frac{m}{\lg m}/2 - n_s k}{3} \times \lceil \delta \lg m \rceil}{m^2}$ flows which can be rewritten as

$$\theta_2 \frac{n_s \sqrt{k}}{\sqrt{n}}$$

for some constant θ_2 when $k = O((\frac{m}{\lg m})^2)$. Then with probability at least $1 - 2e^{-n_s c_8^2/32}$, the minimum data rate that can be supported using any routing strategy, due to the congestion in some quasi-closed cell, is at most

$$O(\frac{(\frac{a}{\sqrt{n}})^{-\beta}\sqrt{n}}{\theta_2 n_s \sqrt{k}}) = O((\frac{a}{\sqrt{n}})^{-\beta} \cdot \frac{1}{n_s} \cdot \frac{\sqrt{n}}{\sqrt{k}}), \quad (15)$$

when the number of receivers k per-flow is $O(\frac{n}{\log^2 n})$. ■