Closing the Gap of Multicast Capacity for Hybrid Wireless Networks

Xiang-Yang Li, Xufei Mao and Shaojie Tang

***Abstract*— We study the *multicast capacity* of a random hybrid wireless network consisting of wireless terminals and base stations. Assume that n wireless terminals (nodes) are randomly deployed in a square region and all nodes have the uniform transmission range r and uniform interference range $R = \Theta(r)$; each wireless node can transmit/receive at W_a-bps. In addition, there are m base stations (neither source nodes nor receiver nodes) that are placed uniformly in this square region; each base station can communicate with adjacent base stations directly with a data rate W_B-bps and the transmission rate between a base station and a wireless node is W_c-bps. Assume that there is a set of n_s randomly selected nodes that will serve as the source nodes of n_s multicast flows (each flow has randomly selected $k-1$ receivers). We found that the multicast capacity for hybrid networks has three regimes and for each of regimes, we derive the matching asymptotic upper and lower bounds of multicast capacity.**

***Index Terms*—Hybrid networks, capacity, multicast, broadcast.**

I. Introduction

The asymptotic capacity of large scale random wireless networks has been widely studied. It is well known that the capacity of a wireless network depends on many aspects of the network, like network architecture, routing strategies, power constraints, interferences and node density, etc. A good understanding of the capacity of different networks will help the users to use current network resources more effectively with respect to different environment and conditions, especially for situations like battlefields, dangerous volcano areas. In pure wireless ad hoc networks, wireless nodes may cooperate in routing each others' packets. However, 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. Another well known and studied network is cellular networks, in which all wireless nodes communicate with the base stations within one hop. In addition, the infrastructure of base stations in cellular network is relative stable, thus guarantee a high performance. However, in some cases, to deploy base stations in some area is not only expensive, but almost impossible.

In this paper, we study the *multicast capacity* of hybrid networks when we choose the *best* protocols for all layers. A *hybrid wireless network* consists of two types of network devices: *base stations* and *wireless terminals*. We assume that all m base stations are regularly placed as a grid in a square region with side-length a meters, and each base station is connected with adjacent base stations by wired lines or wireless channels. Here two base stations are said to be adjacent if their Voronoi regions share a common boundary segment. Assume that each link that connects two base stations has rate W_B-bps and a base station is neither a data source nor a data receiver; it simply serves as a relay gateway. Further we assume that n wireless terminals (with common communication range r and interference range $R = \Theta(r)$) are randomly placed in this square region. When the communication is successful, the data rate between two wireless terminals is W_a-bps. The data rate between a base station and any wireless terminal is W_c-bps. Given all base stations Z, the Voronoi region, denoted as $\mathrm{Vor}(z_i, Z)$, of a base station z_i is called the service *cell* of base station z_i.

We study the capacity of a given hybrid network, and how the capacity of hybrid networks scale with the number of nodes, or with the number of base stations in the networks when a fixed deployment region is given. For most results presented in this paper, we assume that the numbers a, r, n, and m are selected such that the resulting hybrid network is connected with high probability. Due to spatial separation, several wireless nodes can transmit simultaneously provided that these transmissions will not cause *destructive* wireless interferences to any of other transmissions, and all transmissions between two base stations are considered wired links, thus there are no interference to other simultaneous transmissions. Notice that the transmission between a wireless node and a base station is wireless transmission, thus bears interference constraints.

For all randomly distributed n nodes, each node v_i has a randomly chosen $k-1$ destination nodes from other $n-1$ wireless nodes, to which it wishes to send data at an arbitrary data rate λ_i. The minimum per-flow multicast capacity of a random network is defined as $\min_{i=1}^{n_s} \lambda_i$ when there is a schedule of transmissions such that all multicast flows will be received by their destination nodes successfully. For presentation simplicity, we assume that there is only one channel in the wireless networks. As always, we assume that a wireless node has enough memory to buffer all the packets it generates or relay for others such that no packets will be lost through one- or multi-hop transmission. 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 large for some results.

Basically, for multicast, there are three different routing strategies in a hybrid network. The first one is named *Ad Hoc Routing*: given the source node and $k-1$ receivers, a multicast tree using only wireless terminal is constructed and the routing is performed on this tree. This approach has the same capacity as ad hoc wireless network. The second one is based on service cell. For a multicast flow, for each cell that contains at least one receiver (or source node) inside, we construct a tree that spans the receivers (or source node) including the base station

Department of Computer Science, Illinois Institute of Technology, Chicago, IL, USA (e-mail: xli@cs.iit.edu; xmao3@iit.edu; stang7@iit.edu).

in this cell. Then the forest (composed of trees built for each cell) will be connected by links among base stations. This is similar to the routing in cellular networks. We call this routing strategy as *Cellular Routing*. The third routing method can use any subgraph of the original communication graph that spans the receivers and the source node for routing. We call this routing strategy as *Hybrid Routing*. Thus, hybrid networks actually present a tradeoff among traditional BS-oriented network and ad hoc wireless network.

Compared with the similar work by Mao *et al.*, [19], we study more general cases rather than only studying *Cellular Routing* strategy and further close the gap between the upper bound and lower bound in multicast capacity for hybrid networks. Surprisingly, our results show that we can either use *Ad Hoc Routing* strategy or *Cellular Routing* strategy to beat any other routing strategy asymptotically.

The multicast capacity of hybrid wireless networks has been studied in [19]. They assume that all links (links between base stations, links between base stations and ordinary nodes, and links between ordinary nodes) have the same capacity W-bps. They derive asymptotic upper bounds and lower bounds on multicast capacity of the hybrid wireless networks. The total multicast capacity is $O(\frac{\sqrt{n}}{\sqrt{\log n}} \cdot \frac{\sqrt{m}}{k} \cdot W)$ when $k = O(\frac{n}{\log n})$, $k = O(m)$, $\frac{k}{\sqrt{m}} \to \infty$ and $m = o(\frac{a^2}{r^2})$; the total multicast capacity is $\Theta(\frac{\sqrt{n}}{\sqrt{\log n}} \cdot \frac{W}{\sqrt{k}})$ when $k = O(\frac{n}{\log n})$, $k = \Omega(m)$ and $\frac{m}{k} \to 0$. When $k = O(\frac{n}{\log n})$ and $k = O(\sqrt{m})$, the upper bound for the minimum multicast capacity is at most $O(\frac{r \cdot n}{a} \cdot \sqrt{m} \cdot \frac{W}{k})$ and is at least $\Omega(W)$ respectively. When $k = \Omega(\frac{n}{\log n})$, the multicast capacity is $\Theta(W)$. Compared with their results, we assume a more general heterogeneous link capacities and close the gap between lower bounds and upper bounds for all possible cases.

Our Main Contributions: In this paper we derive matching upper bounds and lower bounds on multicast capacity of a hybrid wireless network, in which base stations are distributed regularly in a grid illustrated by Figure 2. Assume that the deployment region and the transmission range r are selected such that the network is connected *w.h.p.* In other words, $n\pi r^2 = \Theta(\log n)$ [13]. We always assume that $m = O(a^2/r^2)$. We show that

Theorem 1: The asymptotic per-flow capacity of n_s multicast sessions by *Cellular Routing* is

$$\vartheta_k(n) = \begin{cases} \Theta(\min(\frac{W_B\sqrt{m}}{n_s\sqrt{k}}, \frac{W_c m}{n_s k}, \frac{W_a m}{n_s k})) & \text{if } k = O(m) \\ \Theta(\min(\frac{W_B}{n_s}, \frac{W_c}{n_s}, \frac{W_a}{n_s})) & \text{if } k = \Omega(m) \end{cases} \quad (1)$$

When the *Ad Hoc Routing* strategy is used, it was proved in [12] that the minimum per-flow multicast capacity is

$$\lambda_k(n) = \begin{cases} \Theta(\frac{a}{r} \cdot \frac{W_a}{n_s\sqrt{k}}) & \text{if } k = O(\frac{a^2}{r^2}) \\ \Theta(\frac{W_a}{n_s}) & \text{if } k = \Omega(\frac{a^2}{r^2}) \end{cases} \quad (2)$$

We then proved that the *Hybrid Routing* strategy will achieve a network capacity at most the larger one of the asymptotic capacity achieved by *Cellular Routing* strategy and the asymptotic capacity achieved by the *Ad Hoc Routing* strategy. Combining the preceding results, we further prove that

Theorem 2: The minimum per-flow capacity $\varphi_k(n)$ by *Hybrid Routing* strategy when $m = O(\frac{a^2}{r^2})$ is of order

$$\begin{cases} \Theta(\max\left[\min\left(\frac{W_B\sqrt{m}}{n_s\sqrt{k}}, \frac{W_c m}{n_s k}, \frac{W_a m}{n_s k}\right), \frac{W_a}{n_s\sqrt{k}}\frac{a}{r}\right]) & \text{if } k = O(m) \\ \Theta(\frac{a}{r} \cdot \frac{W_a}{n_s\sqrt{k}}) & \text{if } k = \Omega(m), k = O(\frac{a^2}{r^2}) \\ \Theta(\frac{W_a}{n_s}) & \text{if } k = \Omega(\frac{a^2}{r^2}) \end{cases} \quad (3)$$

The multicast capacity of hybrid networks when using *Cellular routing* and *Ad Hoc Routing* is shown in Fig. 1. An important finding of our paper is that we prove that the asymptotic multicast capacity achieved in hybrid networks is the upper envelop of those two curves (*Cellular routing* and *Ad Hoc Routing*). So we can either use *Cellular routing* or *Ad Hoc Routing* to "beat" any other routing strategy asymptotically.

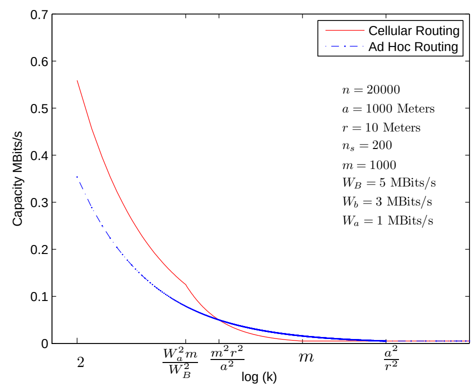


Fig. 1. The capacity bounds (curves) for *Cellular Routing* and *Ad Hoc Routing*. The upper envelop of two curves is the capacity bound for *Hybrid Routing*.

Note that when the transmission range r is smaller, the achievable asymptotic capacity will be larger. However, on the other hand, the transmission range r should be at least a certain value such that the network formed by base stations and terminals will be connected *w.h.p.* It has been shown in [19] that when $\frac{a}{r} \leq \sqrt{\frac{cn\pi}{\log(c\frac{n}{m})+\beta}}$ for $\beta \to \infty$, the resulting network $G = (V \cup Z, E)$ is connected with probability at least $\frac{1}{e^{e^{-\beta}}} \to 1$, when $\beta \to \infty$ and c is constant.

The rest of the paper is organized as follows. In Section II we discuss in detail the network model used in this paper. In Section III and IV we present the matching upper bounds and lower bounds for multicast capacity respectively for the hybrid networks when *Cellular Routing* strategy is used. In Section V, we give the multicast capacity bound when *Hybrid Routing* strategy is used. We review the related results on network capacities in Section VI and conclude the paper in Section VII with the discussion of some possible future works.

II. Network Model

We assume that there is a set $V = \{v_1, v_2, \cdots, v_n\}$ of n *ordinary* wireless terminals randomly deployed in a square region with a side-length a. Each wireless node has transmission range r such that nodes v_i and v_j can communicate successfully iff the Euclidean distance $|v_i - v_j| \leq r$. The data rate of every link v_iv_j is W_a-bps when no interference occurs. A communication from v_i to v_j is interference-free if there is no any other node u that is transmitting and within distance R of receiving node v_j.

We further assume that there are m base stations $Z = \{z_1, z_2, \cdots, z_m\}$ regularly placed in the region. For example, the base stations are placed regularly at positions $(\frac{a}{2\sqrt{m}} + i\frac{a}{\sqrt{m}}, \frac{a}{2\sqrt{m}} + j\frac{a}{\sqrt{m}})$ with $0 \le i \le \sqrt{m}-1$, and $0 \le j \le \sqrt{m}-1$. We generally assume that m is a square of some integer. Fig. 2(a) illustrates a simple example of hybrid networks. Clearly, these m regularly placed base stations divide the original square region into m **cells** as Voronoi diagrams with same side length $\frac{a}{\sqrt{m}}$. We use S_i to denote the cell defined by base station z_i, and for simplicity, by abusing the notation little bit, we say the cell S_i is the service region of base station z_i, *i.e.*, z_i serves as a functional gateway for all wireless nodes in cell S_i when *Cellular Routing* strategy is used. The transmission range of a base station is also assumed to be r. In other words, a base station can only directly serve nodes within distance r. The total data rate that a base station can serve all ordinary wireless nodes is at most W_c-bps with $W_c \ge W_a$. In other words, a base station can serve at most $\frac{W_c}{\lambda}$ flows if each flow requires a data rate λ.

Each base station is connected to its adjacent base stations (at most 4) by wired lines or wireless channels (using frequency different from the frequency used between ordinary wireless nodes). The links between base stations have a large capacity W_B to support traffics. We further assume that $m = o(\frac{a^2}{r^2})$ throughout the paper due to the following observation: when the number of base stations $m \ge \frac{a^2}{r^2}$, all these regularly distributed base stations will cover the whole square, thus a hybrid network will act as a cellular network.

An ordinary node and a base station can communicate with each other only if the Euclidean distance between them is at most r. In other words, the wireless communication range of any base station is also assumed to be r. The complete communication network is a graph $G = (V \cup Z, E)$, where $V = \{v_1, v_2, \cdots, v_n\}$ is the set of ordinary wireless nodes and $Z = \{z_1, z_2, \cdots, z_m\}$ is the set of base stations, and $E = E_a \cup E_B \cup E_c$ is the set of all possible communication links

1) E_a is the set of ad hoc links uv where $u \in V$, $v \in V$, and $\|u - v\| \le r$. Each link in E_a has data rate W_a-bps.
2) E_B is the set of backbone links $z_i z_j$ where $z_i \in Z$, $z_j \in Z$, and $\|z_i - z_j\| = \frac{a}{\sqrt{m}}$. The data rate of each link in E_B is W_B-bps.
3) E_c is the set of cellular links $z_i v_j$ where $z_i \in Z$, $v_j \in V$, and $\|z_i - v_j\| \le r$. The data rate (both up-link and down-link) of each link in E_c is W_c-bps.

For simplicity, we use E_d to denote the set of crossing ad hoc links: $E_d = \{(v_i, v_j) \mid v_i \text{ and } v_j \text{ are from different cells}\}$. We assume that $W_a \le W_c \le W_B$. Given a multicast flow with source v_i and the set of receivers U_i, the routing structure must be a subgraph of G. Three different routing strategies that will be studied here can be categorized as follows

1) *Ad Hoc Routing* strategy will use only the links in E_a. We use $\lambda_k(n)$ to denote the asymptotic multicast capacity achievable by ad hoc routing strategy.
2) *Cellular Routing* strategy will *not* use links in E_d, *i.e.*, $uv \in E_a$ such that u and v are from different cells. We use $\vartheta_k(n)$ to denote the asymptotic multicast capacity achievable by cellular routing strategy.
3) *Hybrid Routing* strategy can use any links in G. We use $\varphi_k(n)$ to denote the asymptotic multicast capacity achievable by hybrid routing strategy.

Please see Fig 2 for illustration. In this paper, we mainly assume that the transmission range r is fixed and thus normalized to one unit throughout the paper.

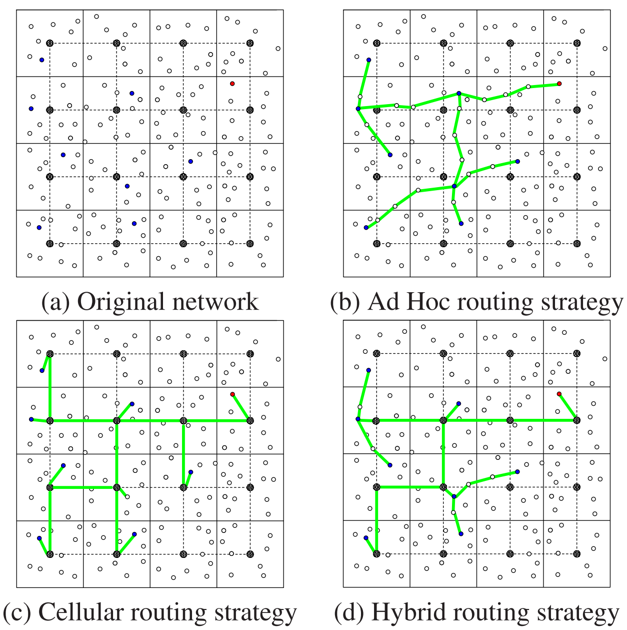


Fig. 2. Illustration of Three Routing Strategies, we use the red node to denote a source node and blues nodes to denote its k receivers

Random Multicast Flows: In this paper, we will concentrate on the *multicast capacity* of a random hybrid network, which generalizes both the unicast capacity [6] and broadcast capacity [9, 14] for random networks (when $m = 0$). Assume that a subset $\mathcal{S} \subseteq V$ of $n_s = |\mathcal{S}|$ random nodes will serve as the source nodes of n_s multicast sessions. We randomly and independently choose n_s multicast sessions. To generate the i-th ($1 \le i \le n_s$) multicast session, k points $p_{i,j} (1 \le j \le k)$ are randomly and independently chosen from the deployment region. Let $v_{i,j}$ be the nearest wireless node from $p_{i,j}$ (ties are broken randomly). Observe that doing this, it is possible that some nodes will serve as a receiver of multiple multicast flows, and a multicast flow may have less then $k-1$ receivers. It is not difficult to show that with high probability, each flow will have at least $(1-\epsilon)(k-1)$ receivers for a small value $0 < \epsilon < 1$. Thus, for simplicity, we always assume that each flow has $k-1$ receivers. In the i-th multicast session, $v_{i,1}$ will be chosen as source node and multicast data to $k-1$ nodes $U_i = \{v_{i,j} \mid 2 \le j \le k\}$ at an arbitrary data rate λ_i.

In this paper, we mainly focus on the protocol interference model induced in [6]. We assume that each node v_i has a fixed interference range R which is within a small constant factor of the transmission range r, *i.e.*, $\varrho_1 r \le R \le \varrho_2 r$ for some constants $1 < \varrho_1 \le \varrho_2$. Under the protocol interference model, any node v_j will be interfered by the signal from v_k if $\|v_k - v_j\| \le R$ where node v_k is sending signal to some node other than v_j.

Capacity Definition: We assume that any node v_i could serve as the source node for some multicast, here $1 \leq i \leq n$. And for each source node v_i, we randomly select $k-1$ receiver nodes from other $n-1$ nodes, say $U_i \subseteq V - \{v_i\}$. Assume that node v_i will send data to these receivers U_i 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 $G = (V \cup Z, E)$, where the node positions of all nodes V, the position of all base stations Z, 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,

Definition 1: Given a network, a multicast rate vector $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_{n-1}, \lambda_n)$ 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 can send $T \cdot \lambda_i$ bits to its corresponding $k-1$ receivers.

The average per flow multicast throughput capacity is defined as $\alpha_k(n) = \frac{\sum_{i=1}^{n} \lambda_i}{n_s}$, where n_s is the number of multicast sessions, and k is the total number of nodes in each multicast session, including the source node. Similarly, given n_s multicast sessions with S as source nodes, the minimum per-flow multicast capacity is defined as

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

In this paper, we will focus on the minimum per-flow capacity.

Definition 2: 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

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

Throughout this paper, we will focus on studying the minimum per-flow multicast capacity which is defined as $\varphi_k(n) = \min_{v_i \in S} \lambda_i$.

III. Upper Bounds in Multicast Capacity by Cellular Routing

When *Cellular Routing* is used, the capacity for a hybrid network can be constrained due to three different congestion scenarios: (1) the backbone formed by the links E_B is congested; (2) the cellular links E_c are congested; and (3) the ad hoc links $E_a \setminus E_d$ in some cell are congested. We will derive an upper bound separately on minimum per-flow multicast capacity for each of the aforementioned three conditions.

Technique Lemmas: Throughout this paper, we will repeatedly use these lemmas.

Lemma 3: For the i-th flow, let $k_{i,j}$ be the number of terminals that will fall inside the service cell of the jth base station z_j. Then, $k_{i,j}$ is a random variable with mean $E(k_{i,j}) = \frac{k}{m}$ and variance $\mathrm{Var}(k_{i,j}) = \frac{k}{m}(1 - \frac{1}{m})$.

Note that $\mathbf{Pr}\left(k_{i,j} = t\right)$ is $\binom{k}{t}\left(\frac{1}{m}\right)^t\left(1 - \frac{1}{m}\right)^{k-t}$.

Lemma 4: Let variable $k'_{i,j}$ denotes the number of terminals of the i-th flow that fall inside the cell of the base station z_j, but not inside the communication disk of z_j (centered at z_j with radius r). Then

$$\mathbf{Pr}\left(k'_{i,j} = t\right) = \binom{k}{t}\left(\frac{1}{m} - \frac{r^2}{a^2}\right)^t\left(1 - \frac{1}{m} + \frac{r^2}{a^2}\right)^{k-t}$$

Then its mean is $E(k'_{i,j}) = k(\frac{1}{m} - \frac{r^2}{a^2})$ and variance $\mathrm{Var}(k'_{i,j}) = k(\frac{1}{m} - \frac{r^2}{a^2})(1 - \frac{1}{m} + \frac{r^2}{a^2})$. Recall that in this paper, we assumed that the number of base stations $m \leq c\frac{a^2}{r^2}$ for some constant $0 < c < 1$. Thus, $\frac{1}{m} - \frac{c^2}{a^2} \geq (1-c)\frac{1}{m}$. This implies that $E(k'_{i,j}) \geq (1-c)\frac{k}{m}$ and variance $(1-c)k\frac{1}{m}(1-(1-c))\frac{1}{m}) \leq \mathrm{Var}(k'_{i,j}) \leq k\frac{1}{m}(1-\frac{1}{m})$.

Lemma 5: Let variable $X_{i,j} \in \{0, 1\}$ denote whether the j-cell (defined by base station z_j) contains some terminals from the ith flow, *i.e.*, $X_{i,j} = 1$ if $k_{i,j} > 0$, and $X_{i,j} = 0$ if $k_{i,j} = 0$. Thus, $\mathbf{Pr}\left(X_{i,j} = 1\right) = 1 - (1 - \frac{1}{m})^k$. In addition, $\mathrm{Var}(X_{i,j}) = E(X_{i,j}^2) - E(X_{i,j})^2 = (1 - (1 - \frac{1}{m})^k)(1 - \frac{1}{m})^k$.

Lemma 6: Let variable f_j denote the number of flows, each of which has at least a terminal node inside the j-th cell. Then $f_j = \sum_{i=1}^{n_s} X_{i,j}$. In addition, $E(f_j) = n_s(1 - (1 - \frac{1}{m})^k)$ and variance $\mathrm{Var}(f_j) = n_s \cdot \mathrm{Var}(X_{i,j}) = n_s(1 - (1 - \frac{1}{m})^k)(1 - \frac{1}{m})^k$.

Lemma 7: Let variable $X'_{i,j} \in \{0, 1\}$ denote whether some terminals from the i^{th} flow fall into the j^{th}-cell, but not inside the communication disk centered at z_i, $X'_{i,j} = 1$ if $k'_{i,j} > 0$, and $X'_{i,j} = 0$ if $k'_{i,j} = 0$. Thus, $\mathbf{Pr}\left(X'_{i,j} = 1\right) = 1 - (1 - (\frac{1}{m} - \frac{r^2}{a^2}))^k$. In addition, $\mathrm{Var}(X'_{i,j}) = E(X_{i,j}^2) - E(X_{i,j})^2 = (1 - (1 - (\frac{1}{m} - \frac{r^2}{a^2}))^k)(1 - (\frac{1}{m} - \frac{r^2}{a^2}))^k$.

Lemma 8: Let variable f'_j denote the number of flows, each of which has at least a terminal node inside the j-th cell, but not inside the communication disk centered at z_j. Then $f'_j = \sum_{i=1}^{n_s} X'_{i,j}$. In addition, $E(f'_j) = n_s(1 - (1 - (\frac{1}{m} - \frac{r^2}{a^2}))^k)$ and variance $\mathrm{Var}(f'_j) = n_s \cdot \mathrm{Var}(X'_{i,j}) = n_s(1 - (1 - (\frac{1}{m} - \frac{r^2}{a^2}))^k)(1 - (\frac{1}{m} - \frac{r^2}{a^2}))^k$.

Lemma 9: Let variable k_i denote the number of cells that has at least one terminal from flow i inside. Clearly, $k_i = \sum_{j=1}^{m} X_{i,j}$. Then $E(k_i) = m(1 - (1 - \frac{1}{m})^k)$ and variance $\mathrm{Var}(k_i) = m \cdot \mathrm{Var}(X_{i,j}) = m(1 - (1 - \frac{1}{m})^k)(1 - \frac{1}{m})^k$. Furthermore, when $k \leq m$, $\frac{k}{2} \leq E(k_i) \leq \min(m, 2k)$. When $k > m$, $\frac{m}{2} \leq E(k_i)$.

Given a routing strategy $\mathcal{A}$, let $T_i(\mathcal{A})$ be the tree used to route the i-th flow. When $\mathcal{A}$ is clear from the context, we will simplify it as T_i by dropping $\mathcal{A}$. The following lemma was shown in [22].

Lemma 10: Given k points Q randomly placed in a square of side length a, the Euclidean minimum spanning tree, denoted as $\mathrm{EMST}(Q)$, has an expected total edge length $\Theta(\sqrt{k}a)$ and its variance $\mathrm{Var}(\mathrm{EMST}(Q)) \ll a^2 \cdot \log k$.

It was proved in [11, 12] that any routing tree T_i for a set Q of random k points in the square of side-length a, its total edge length is at least $\frac{1}{2}$ times the total edge length of $\mathrm{EMST}(Q)$.

A. Upper Bound Due to Links in E_B

The upper bound on multicast capacity due to links in E_B has two regimes: $k = O(m)$ and $k = \Omega(m)$.

1) When $k = O(m)$: In this case, for each flow i, we let B_i be the set of base stations whose service cell contains at least one terminal from the i-th flow. Then we need build a connected structure using only links in E_B to span B_i. Let T_B^i be the tree (covering all base stations in B_i) constructed by a given routing method. Then we know that $|T_B^i| \geq |\mathrm{EMST}(B_i)|/2$. Hereafter, if $\mathcal{S}$ is a set, we use $|\mathcal{S}|$ to denote the cardinality of S; if $\mathcal{S}$ is a tree, we use $|\mathcal{S}|$ to denote the total Euclidean length of tree $\mathcal{S}$. Notice that the set B_i is a random variable and $|B_i| = k_i$, where random variable k_i is as defined before. Similar to [22], we can prove the following lemma:

Lemma 11: Given k_i base stations B_i randomly selected and all base stations are placed in a square region of side-length a, the Euclidean minimum spanning tree $\mathrm{EMST}(B_i)$ has an expected total edge length $c_1\sqrt{k_i}a$ for a constant $c_1 \in (0, 2\sqrt{2}]$ and its variance $\mathrm{Var}(|\mathrm{EMST}(B_i)|) \ll a^2 \cdot \log k_i$.

Theorem 12: When $k \leq \theta_0 m$ for some constant θ_0, there is a constant c_3 such that, with probability at least $1 - 2e^{-n_s/8}$, the minimum data rate that can be supported using cellular routing strategy is at most $\frac{W_B\sqrt{m}}{c_3 n_s\sqrt{k}}$ for any routing strategy due to the congestion in backbone links.

Proof: Let $C(T_B^i)$ denote the number of cells that the routing tree T_B^i will use, *i.e.*, the number of base stations used in T_B^i. Obviously, $C(T_B^i) \geq k_i$, the number of cells that contain the receivers of the i-th flow. Notice that each base station is connected to at most 4 adjacent base stations. Then $|T_B^i|/(4\frac{a}{\sqrt{m}}) \leq C(T_B^i) = |T_B^i|/(\frac{a}{\sqrt{m}})$. Let variable $L = \sum_{i=1}^{n_s} C(T_B^i)$, denoting the total load of all cells. Here the load of a cell by a routing method is the number of flows passing the cell for the multicast tree constructed. Then $L \geq \sum_{i=1}^{n_s} |T_B^i|/(4\frac{a}{\sqrt{m}}) \geq \sum_{i=1}^{n_s} |\mathrm{EMST}(B_i)|/(8\frac{a}{\sqrt{m}})$. Notice that $E(\sum_{i=1}^{n_s} |\mathrm{EMST}(B_i)|) = n_s c_1 E(\sqrt{k_i})a$ and $\mathrm{Var}(\sum_{i=1}^{n_s} |\mathrm{EMST}(B_i)|) \ll n_s^2 a^2 \log k_i$. Thus $E(L) \geq c_1 n_s E(\sqrt{k_i})\sqrt{m}/8$.

We then compute the value $E(\sqrt{k_i})$. Recall that variable $X_{i,j}$ denotes whether the j-th cell contains any terminal from the i-th flow and $k_i = \sum_{j=1}^m X_{i,j}$. By definition, $E(\sqrt{k_i}) = E(\sqrt{\sum_{j=1}^m \mathbf{Pr}(X_{i,j} = 1)}) = \sqrt{m(1 - (1 - \frac{1}{m})^k)}$. Then,

$$\sqrt{\min(m, k)/2} \leq E(\sqrt{k_i}) \leq \sqrt{\min(m, 2k)}$$

When $k \leq \theta_0 m$, we have $E(\sqrt{k_i}) \geq c_2\sqrt{k}$ for a constant $c_2 = \sqrt{\min(\frac{1}{2\theta_0}, 1)}$.

Define random variables $X_q = \sum_{j=1}^q (|\mathrm{EMST}(B_j)| - E(|\mathrm{EMST}(B_j)|))$. Then $E(X_{q+1} \mid X_1, \cdots, X_q) = X_q$, *i.e.*, variables X_i are martingale. In addition, $|X_q - X_{q-1}| = ||\mathrm{EMST}(B_q)| - E(|\mathrm{EMST}(B_q)|)| \leq |\mathrm{EMST}(B_q)|$, which is $\leq 2\sqrt{2}\sqrt{k_i}a \leq 2\sqrt{2}\sqrt{k}a$. From Azuma's Inequality, we have $\mathbf{Pr}(|X_{n_s} - X_0| \geq t) \leq 2\exp(-\frac{t^2}{2\sum_{i=1}^{n_s} 8ka^2})$.

Let $t = \epsilon \sum_{i=1}^{n_s} E(|\mathrm{EMST}(B_i)|)$. Clearly, $\epsilon n_s c_1 c_2\sqrt{k}a \leq t \leq 2\sqrt{2}\epsilon\sqrt{k}a$. Note that $X_0 = 0$. Then,

$$\begin{aligned} &\mathbf{Pr}\left(\sum_{i=1}^{n_s} |\mathrm{EMST}(B_i)| \leq \sum_{i=1}^{n_s} E(|\mathrm{EMST}(B_i)|) - t\right) \\ \leq\ &\mathbf{Pr}(|X_{n_s}| \geq t) \leq \exp(-\frac{t^2}{2\sum_{i=1}^{n_s} 8ka^2}) \\ \leq\ &\exp(-\frac{(\epsilon n_s c_1 c_2\sqrt{k}a)^2}{8n_s ka^2}) = \exp(-\frac{n_s\epsilon^2 c_1^2 c_2^2}{8}) \end{aligned}$$

Consequently, for a constant $\epsilon \in (0, 1)$, we have

$$\mathbf{Pr}\left(\sum_{i=1}^{n_s} |\mathrm{EMST}(B_i)| \leq (1 - \epsilon)n_s c_1 E(\sqrt{k_i})a\right) \leq 2e^{-\frac{n_s\epsilon^2 c_1^2 c_2^1}{8}},$$

$$\mathbf{Pr}\left(\sum_{i=1}^{n_s} |\mathrm{EMST}(B_i)| \geq \sum_{i=1}^{n_s} c_1 E(\sqrt{k_i})a/2\right) \geq 1 - 2e^{-n_s c_1^2 c_2^2/32}.$$

Then,

$$\mathbf{Pr}\left(L \geq n_s c_1 E(\sqrt{mk_i})/16\right) \geq 1 - 2e^{-n_s c_1^2 c_2^2/32}.$$

It implies

$$\mathbf{Pr}\left(L \geq n_s c_1 c_2\sqrt{km}/16\right) \geq 1 - 2e^{-n_s c_1^2 c_2^2/32} \text{ if } k \leq \theta_0 m.$$

Recall that L denotes the total load of all cells. Then by Pigeonhole principle, with probability at least $1 - 2e^{-n_s c_1^2 c_2^2/32}$, there is at least one cell, that will be used by at least $\frac{n_s c_1 c_2\sqrt{km}}{m}$ flows. Thus, with probability at least $1 - 2e^{-n_s c_1^2 c_2^2/32}$, the minimum data rate that can be supported using cellular routing strategy is at most $\frac{W_B}{\frac{n_s c_1 c_2\sqrt{km}}{m}} = \frac{W_B\sqrt{m}}{c_1 c_2 n_s\sqrt{k}}$ for any routing strategy due to the congestion in backbone links. By letting $c_3 = c_1 c_2$ finishes the proof of the theorem. ■

2) When $k = \Omega(m)$: Recall that in this case, we have shown that $E(k_i) \geq m/2$, *i.e.*, for each flow, the expected number of cells that will contain its terminals is at least $m/2$. More precisely, it is easy to show that, for any cell j, the probability, $\mathbf{Pr}(X_{i,j} = 1)$, that it will contain a terminal from flow i is at least $1 - 1/e > 1/2$. Then using Azuma's Inequality, we can prove that, with probability at least $1 - 2e^{-n_s/8}$, the total load $L \geq n_s m/4$. Thus, by Pigeonhole principle, there is one cell such that its load (the number of flows using its base-station) is at least $n_s/4$. Consequently, we have the following theorem.

Theorem 13: When $k \geq \theta_0 m$ for some constant $\theta_0 > 1$, with probability at least $1 - 2e^{-n_s/8}$, the minimum data rate that can be supported using cellular routing strategy is at most $\frac{4W_B}{n_s}$ for any routing strategy due to the congestion in backbone.

B. Upper Bound Due to Links in E_c

In this subsection, we study the minimum per-flow data rate due to the congestion when ordinary wireless nodes access the base-stations in their cells. Recall that we assume that both the uplink rate and the down-link rate between the base-station and the ordinary wireless nodes in its cell is W_c-bps. We will study upper bounds based on two subcases, whether $k = O(m)$ or not. We essentially will study the number of flows f_j inside of j^{th} cell that will pass through a base-station z_j.

1) When $k = O(m)$: We first study the case when the number of terminals per-flow is $k = O(m)$. Notice that when $k \le m$, $E(f_j) = n_s(1-(1-\frac{1}{m})^k) > \frac{k}{2m}n_s$ and $\mathrm{Var}(f_j) < \frac{2k}{m}n_s$.

Lemma 14: When n_s satisfies the condition (4), the variable $\max_{j=1}^{m} f_j$ is $\Theta(n_s\frac{k}{m})$ with probability at least $1-\frac{1}{n}$.

Proof: We use the VC-Theorem to prove this lemma. Let the set $\mathcal{C} = \{\mathrm{Vor}(z_j, Z) \mid 1 \le j \le m\}$ be the class of cells defined by all base-stations. Let F_i be the i-th flow and F_i is said to "belong to" the j-th cell if some of its terminals is contained inside the j-th cell $\mathrm{Vor}(z_j, Z)$, which is denoted as $F_i \in \mathrm{Vor}(z_j, Z)$. Then $f_j = \sum_{F_i} I(F_i \in \mathrm{Vor}(z_j, Z))$, where $I(F_i \in \mathrm{Vor}(z_j, Z)) = 1$ if $F_i \in \mathrm{Vor}(z_j, Z)$ and $I(F_i \in \mathrm{Vor}(z_j, Z)) = 0$ otherwise. Obviously, $\text{VC-d}(\mathcal{C}) \le \log m$ since the cardinality of $\mathcal{C}$ is m. In addition, the probability $P(A)$ that a flow "belongs to" a cell A is $P(A) = 1-(1-\frac{1}{m})^k$. It is easy to show that, when $0 < k < m$, we have $\frac{k}{2m} < P(A) < \frac{2k}{m}$. Then by VC-Theorem, we know that for every $\epsilon, \delta > 0$, $\mathbf{Pr}\left(\sup_{A\in\mathcal{C}}\left|\frac{\sum_{i=1}^{n_s} I(F_i\in A)}{n_s} - P(A)\right| \le \epsilon\right) > 1-\delta$ whenever $n_s > \max\left\{\frac{8\cdot\text{VC-d}(\mathcal{C})}{\epsilon}\cdot\log\frac{13}{\epsilon}, \frac{4}{\epsilon}\log\frac{2}{\delta}\right\}$. When we choose the parameters $\epsilon = \frac{k}{4m}$, $\delta = \frac{1}{n}$, and

$$n_s > \max\left(\frac{32m\log m}{k}\log\frac{52m}{k}, \frac{16m}{k}\log(2n)\right), \quad (4)$$

Then

$$\mathbf{Pr}\left(\sup_{i=1}^{m}|f_i - n_sP(A)| \le n_s\frac{k}{4m}\right) > 1-\frac{1}{n}.$$

Hence, $\mathbf{Pr}\left(\forall i \in [1,m], n_s\frac{k}{4m} \le f_i \le n_s\frac{9k}{4m}\right) > 1-\frac{1}{n}$. ■

Based on the preceding lemma, we conclude that,

Theorem 15: When $k \le m$, the rate due to the congestion of accessing the base-stations, with probability at least $1-\frac{1}{n}$, is at most $\frac{W_c}{\max_i f_i} \le \frac{W_c\cdot(4m)}{n_s k} = O(\frac{W_c\cdot m}{n_s k})$.

2) When $k = \Omega(m)$: We then study an upperbound on the rate achievable due to the congestion of accessing base-stations when $k > m$. In this case, $n_s > E(f_j) = n_s(1-(1-\frac{1}{m})^k) > n_s(1-\frac{1}{e})$.

Lemma 16: When n_s satisfies the condition (5), the variable $\max_{j=1}^{m} f_j$ is also $\Theta(n_s)$ with probability at least $1-\frac{1}{n}$.

Proof: Similar as the proof for Lemma 14, except that when $k \ge m$, we have $1-\frac{1}{e} < P(A) < 1$. Based on VC-Theorem, by choosing the parameters $\epsilon = \frac{1}{e}$, $\delta = \frac{1}{n}$, we know that when

$$n_s > \max\left(8e\log m\log(13e), 4e\log(2n)\right), \quad (5)$$

Then

$$\mathbf{Pr}\left(\sup_{i=1}^{m}|f_i - n_sP(A)| \le n_s\frac{1}{e}\right) > 1-\frac{1}{n}.$$

Hence, we have

$$\mathbf{Pr}\left(\forall i \in [1,m], n_s(1-\frac{2}{e}) \le f_i \le n_s\right) > 1-\frac{1}{n}.$$

This finishes the proof. ■

Obviously, we have the following theorem.

Theorem 17: When $k \ge m$, with probability at least $1-\frac{1}{n}$, the minimum per-flow rate by any *Cellular Routing* strategy is at most $\frac{W_c}{n_s(1-\frac{2}{e})}$.

C. Upper Bound Due to Links in $E_a \setminus E_d$

In previous subsections, we study upper bounds on the multicast capacity in hybrid networks due to the congestion at the backbone links (connecting pairs of base-stations), and due to the congestion in accessing the base-stations. We now focus on studying the capacity upper bounds due to the congestion in ad hoc links $E_a \setminus E_d$.

A trivial upper bound for total multicast capacity is $W_a \cdot n$ since there are n source nodes in total and each can send data at W_a bits/sec. However, we can make the upper bounds more tight due to the following observations. For each source node v_i, when we multicast the data from one source node v_i to all its $k-1$ receivers in set $U_i = \{v_{i_1}, v_{i_2}, \cdots, v_{i_{k-1}}\}$, the resulting multicast tree will contain at least k nodes, and possibly more. More possibly, when a non-leaf node v in the multicast tree sends data to its children, **all** nodes that are within its transmission range will receive the data or at least they cannot transmit successfully at the same time no matter these nodes are intended receivers or not. In this case, we say all these nodes are *charged* a *copy* of the data. To study the multicast capacity, we partition the deployment square into grids of size r. Clearly, there are at most $\lceil\frac{a}{r}\rceil^2 = \Theta(\frac{a^2}{r^2})$ such grid cells. Notice that among such grid cells, some of them can be directly reached by some base-stations. Let g be the total number of grid cells that is disjoint from the union of disks $\bigcup_{j=1}^{m} D(z_j, r)$. Then obviously $g \ge \lceil\frac{a}{r}\rceil^2 - 9m = \Theta(\frac{a^2}{r^2})$ when $m \le \frac{a^2}{10r^2}$. Here, the constant 9 comes from the fact that any base station only can over at most 9 grids of size r at the same time due to our previous assumption that the transmission range of each base station is also r. Thus, throughout this paper, we assume that $m \le \frac{a^2}{10r^2}$.

Recall that we assume that the interference range $R > \varrho_1 r$. Then at any time instance, the distance between two active senders v_1 and v_2 is at least $R - r \ge (\varrho_1 - 1)r$. Consequently, we have

Lemma 18: For any grid of side-length r, there are at most a constant number (denoted as $\kappa < (1+\frac{2}{\varrho_1-1})^2$) of nodes inside the grid that can send data simultaneously without causing interference to receivers.

This lemma implies that the total data that can be sent out from any grid during any time interval t is at most $W_a\cdot\kappa t$ for a constant κ. To prove an upper bound on the capacity, we will only consider the grid cells that are disjoint from the disks defined by base-stations. In other words, for nodes located inside these grid-cells, it cannot reach the base-stations directly and its data have to be relayed by some other nodes to reach the base stations. Given a routing strategy, for the i-th flow and j-th grid cell, let $Y_{i,j}$ be the variable denote whether the i-th flow will be routed through the j-th grid cell by this routing strategy. Let $Y_j = \sum_{i=1}^{n_s} Y_{i,j}$ the total number of flows that will be routed through the j-th grid cell by this routing strategy. Then from Lemma 18, we can conclude that the minimum per-flow data rate is at most

$$\frac{W_a\cdot\kappa}{\max_{j=1}^{g} Y_j} \quad (6)$$

The rest of the subsection is devoted to give a better lower bound on $\max_{j=1}^{g} Y_j$, thus a tighter upper bound of multicast capacity of the hybrid wireless network. Notice that the bound on $\max_{j=1}^{g} Y_j$ depends on the routing strategy. We actually will prove that, regardless of routing strategies used, $\max_{j=1}^{g} Y_j$ will be at least a certain value *w.h.p*. For simplicity, hereafter when we say k receivers, we mean that one source node pluses all its $k-1$ receivers.

1) When $k = O(m)$: As we know, under the *Cellular Routing* strategy, all flows inside of a cell with side-length $\frac{a}{\sqrt{m}}$ will firstly go to the closest base station by one- or multi-hop. Before the traffic reach the base station, the last hop transmission is a cellular link and the second to the last hop link is an ad hoc link. Thus, the potential congestion will happen on those second to last hop ad hoc links for the following two reasons: a) Each base station only has transmission range r and can cover (touch) a relatively small area (at most 9 adjacent cells with side length r around the base station.) b) Intuitively, the cell closed to the base station will have much burden to relay ad hoc traffic to the base station. Clearly, it is equivalent to study the number of flows f'_j inside of j^{th} cell but not inside of the communication disk centered at z_i.

Lemma 19: When n_s satisfies the condition (7), the variable $\max_{j=1}^{m} f'_j$ is $\Theta(n_s \frac{k}{m})$ with probability at least $1 - \frac{1}{n}$.

Proof: We use the similar proof used in Lemma 14 to prove this. Assume the set $\mathcal{C}' = \{\text{Vor}(z'_j, Z) \mid 1 \le j \le m\}$ be the class of regions (each region is the cell without the communication disk of the base station) defined by all base-stations. The i^{th} flow F'_i is said to "belong to" the j^th cell if some of its terminals is contained inside the j-th cell, but not inside the communication disk centered at z_j, $\text{Vor}(z'_j, Z)$, which is denoted as $F'_i \in \text{Vor}(z'_j, Z)$. In addition, the probability $P(A')$ that a flow "belongs to" a region A' is $P(A') = 1-(1-(\frac{1}{m} - \frac{r^2}{a^2}))^k$. It is easy to show that, when $0 < k < \frac{10m}{9}$ and $m \ge \frac{a^2}{10r^2}$ we have $\frac{9k}{20m} < P(A') < \frac{9k}{5m}$. By the same argument in Lemma 14 and VC-Theorem, we have

$$\forall \epsilon, \delta > 0, \mathbf{Pr}\left(\sup_{A' \in \mathcal{C}'} \left| \frac{\sum_{i=1}^{n_s} I(F'_i \in A')}{n_s} - P(A') \right| \le \epsilon \right) > 1 - \delta,$$

whenever $n_s > \max\left\{\frac{8 \cdot \text{VC-d}(\mathcal{C})}{\epsilon} \cdot \log \frac{13}{\epsilon}, \frac{4}{\epsilon} \log \frac{2}{\delta}\right\}$. When we choose the parameters $\epsilon = \frac{k}{4m}$, $\delta = \frac{1}{n}$, and

$$n_s > \max\left(\frac{32m \log m}{k} \log \frac{52m}{k}, \frac{16m}{k} \log(2n)\right), \quad (7)$$

we have

$$\mathbf{Pr}\left(\sup_{i=1}^{m} |f'_i - n_s P(A')| \le n_s \frac{k}{4m}\right) > 1 - \frac{1}{n}.$$

Hence, $\mathbf{Pr}\left(\forall i \in [1, m], n_s \frac{k}{5m} \le f'_i \le n_s \frac{41k}{20m}\right) > 1 - \frac{1}{n}$. ■

Based on the preceding lemma, we conclude that

Theorem 20: When $k \le \frac{10m}{9}$ and n_s satisfies the condition (7), the rate due to the congestion of ad hoc links, with probability at least $1 - \frac{1}{n}$, is $O(\frac{W_a m}{n_s k})$.

Proof: By Lemma 19, for j^{th} cell, the number of ad hoc flows f'_j which will converge to base station z_j, $\max_{j=1}^{m} f'_j = \theta_2 \frac{n_s k}{m}$ with probability at least $1 - \frac{1}{n}$ where n_s satisfies the condition (7) and θ_2 is some positive constant. In addition, there are at most 9 cells with side-length r to relay these flows such that there exist at least one cell has to relay at least $\frac{\theta_2}{9} \frac{n_s k}{m}$ flows by Pigeonhole principle. Therefore, by equation 6, the min-flow capacity cannot exceed $\frac{W_a \cdot \kappa}{\frac{\theta_2}{9} \frac{n_s k}{m}} = O(\frac{W_a m}{n_s k})$ where κ and θ_2 are constants. ■

2) When $k = \Omega(m)$ and $k = O(\frac{a^2}{r^2})$: We are going to study an upper bound on the rate achievable due to the congestion of all ad hoc flows (links) which targets to or from the base station when $k > m$. Recall that we use f'_j to denote all the ad hoc flows which exist in j^{th} cell and as we have shown before, the expected value of $E(k'_{i,j}) \ge (1-c)\frac{k}{m})$ for some constant c. It is not difficult to show that $n_s > E(f'_j) = n_s(1-(1-(\frac{1}{m} - \frac{r^2}{a^2}))^k) > n_s(1-(\frac{1}{e})^{1-c})$. Next, we show that the maximum number of ad hoc flows inside some cell is the constant fraction of total n_s multicast flows by the following Lemma 21.

Lemma 21: When n_s satisfies the condition (8), the variable $\max_{j=1}^{m} f'_j$ is also $\Theta(n_s)$ with probability at least $1 - \frac{1}{n}$.

Proof: Similar as the proof for Lemma 19 except that when $k \ge m$ we have $1 - (\frac{1}{e})^{1-c} < P(A') < 1$. Based on VC-Theorem, by choosing parameters $\epsilon = (\frac{1}{e})^{1-c}$, $\delta = \frac{1}{n}$, we know that when

$$n_s > \max\left(8e^{1-c} \log m \log(13e^{1-c}), 4e^{1-c} \log(2n)\right), \quad (8)$$

$\mathbf{Pr}\left(\sup_{i=1}^{m} |f'_i - n_s P(A)| \le n_s (\frac{1}{e})^{1-c}\right) > 1 - \frac{1}{n}$. Hence, $\mathbf{Pr}\left(\forall i \in [1, m], n_s(1 - \frac{2}{e^{1-c}}) \le f'_i \le n_s\right) > 1 - \frac{1}{n}$. ■

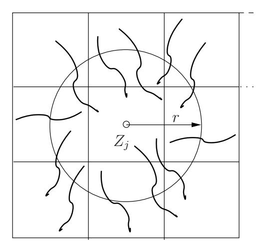


Fig. 3. The scenario where multiple ordinary nodes are close to the base station z_j in the j^{th} grid cell.

Because all the ad hoc flows inside of a cell will finally go to or come from the base station based on the *Cellular Routing* strategy, the potential congestion will happen when the ad hoc flows converge into the communication disk of central base station. See Fig. 3 for illustration. Because the base station z_j has transmission range r, z_j's service range can cover at most 9 adjacent small grids with side length r. In other words, all the ad hoc flows inside the j^{th} cell have to go through these 9 small cells. From Lemma 18, each cell with side length r can only have κ concurrent transmitters. Therefore, by Equation 6 and Lemma 21, we have

Theorem 22: With probability at least $1 - \frac{1}{n}$, the minimum per-flow rate for n_s multicast sessions is bounded by $O(\frac{W_a}{n_s}$ where $k = \Omega(m)$ and $k = O(\frac{a^2}{r^2})$. due to the capacity constraints of ad hoc links in $E_a \setminus E_d$.

3) When $k = \Omega(\frac{a^2}{r^2})$: In the previous subsection, we showed an upper bound of the multicast capacity of hybrid network when $k < \theta_1 \cdot a^2/r^2$. In this subsection we will present an upper bound on multicast capacity when $k \ge \theta_1 \cdot a^2/r^2$.

In [11], Li has proved that when $k = \Omega(\frac{a^2}{r^2})$, the union of the transmission disks of these k receiver nodes in a multicast session will cover at least a constant fraction, say $0 < \rho_2 \leq 1$, of the deployment region. Thus the minimum per-flow capacity of hybrid network due to the congestion of ad hoc link will approximately be equal to the broadcast capacity, *i.e.*, $O(\frac{W_a}{n_s})$. Combined with Theorem 13 and Theorem 17, we have the following theorem:

Theorem 23: When $k \geq \theta \cdot a^2/r^2$ for a constant θ, the minimum per-flow multicast capacity for the hybrid network is bounded by $O(\min(\frac{W_B}{n_s}, \frac{W_c}{n_s}, \frac{W_a}{n_s}))$ with high probability.

IV. Lower Bounds in Multicast Capacity by Cellular Routing

In this section, we will derive asymptotically lower bound in the multicast capacity by presenting a multicast scheme.

A. Implement of Routing Strategies

We proposed the following multicast routing strategy for *Cellular Routing* in Algorithm 1. As we have explained before, based on the *Cellular Routing* strategy, each receiver node will try to reach or be reached by the closest base station by one- or multi-hop. Assume set $U^i = \{v_1^i, v_2^i, \cdots, v_k^i\}$ is the union set of source node v^i and its randomly selected $k-1$ receivers for the i^{th} multicast flow, here we assume $v_k^i = v_i$ for simplicity. Assume U_j^i is the node set containing all receivers of the i^{th} multicast flow which are falling into the j^{th} cell. Obviously, $U^i = \bigcup_{j=1}^{m} U_j^i$. We further assume set $Z^i = \{z_1^i, z_2^i, \cdots, z_t^i\}$ contains all the base stations, each of whose cell contains at least one receiver of the i^{th} multicast flow. Clearly, $1 \leq t \leq k$.

Algorithm 1 *Cellular Routing* strategy for i^{th} multicast flow

Input: U^i

1: Compute Z_i based on U^i, then construct a Minimum Spanning Tree (MST) which contains all nodes in Z_i (may need other base station as internal nodes) by backbone links only. Assume the root of the constructed MST is the base station (say z_s) which falls in the same cell as the source node v_i does. Then do broadcasting from z_s to the other base stations on the MST.
2: **for** each cell S_j which contains at least one receiver inside **do**
3: if S_i contains the source node v_i, then v_i find a shortest path connecting to z_i
4: if S_i contains at least one receiver from $k-1$ receivers, construct a BFS tree from the root z_i which covers all receivers inside. This may need other non-receiver nodes as internal nodes on the BFS tree. Then do broadcasting from base station z_j to all wireless nodes on the constructed BFS tree.
5: **end for**

In the following section, we will analyze the lower bound multicast capacity where $k = O(\frac{a^2}{r^2})$ and $k = \Omega(\frac{a^2}{r^2})$ separately as we did in the previous sections. 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 in each cell S_i which spanning k_i receivers inside and thus obtain a multicast forest which spans all k receivers. Next, we will show the lower bound capacity achievable by the Algorithm 1 under different cases.

B. When $k = O(m)$

When the number of receivers of each multicast session satisfies $k = O(m)$, we analyze the minimum per-flow lower bound capacity achievable by backbone links, cellular links and ad hoc links one by one. For each multicast flow, we use Algorithm 1 to do routing.

We first introduce the lower bound capacity achievable by the backbone links. We know for each multicast flow, the broadcast capacity on the MST tree constructed in Algorithm 1 is $\Theta(W_B)$ due to the result in [9]. In addition, according to the result (Theorem 31) in [11]) we know that if there are n_s random multicast flows in a square region with side-length a, there is a sequence of $\delta(n) \to 0$ such that for any square cell s with side-length $\frac{a}{\sqrt{m}}$ inside of the square region,

Pr $\left(\text{\# of flows using } s \leq \frac{3\delta_3 n_s}{2} \frac{\sqrt{k}\frac{a}{\sqrt{m}}}{a} \right) = \frac{3\delta_3 n_s}{2} \frac{\sqrt{k}}{\sqrt{m}}$ where δ_3 is some constant. Hence, *w.h.p*, the number of flows needed to be relayed by any base station is no more than $\frac{3\delta_3 n_s}{2} \frac{\sqrt{k}}{\sqrt{m}}$. Therefore, the lower bound capacity for backbone links is at least $\Omega(\frac{W_B\sqrt{m}}{n_s\sqrt{k}})$ by Algorithm 1 with a TDMA schedule.

Next, we show the lower bound capacity achievable by cellular links when $k = O(m)$. By Lemma 14, we know for all m cells, when n_s satisfies the condition (4), the variable $\max_{j=1}^{m} f_j$ is $\Theta(n_s \frac{k}{m})$ with probability at least $1 - \frac{1}{n}$. Here, f_j denotes the number of flows inside of j^{th} cell that will pass through the base station z_j. Thus, for any base station, by a simple TDMA schedule, the achievable lower bound capacity for cellular links is $\Omega(\frac{W_c m}{n_s k})$.

The remaining part of this subsection, we show the lower bound capacity achievable by ad hoc links using *Cellular Routing* when $k = O(m)$. Recall that after applying Algorithm 1, each multicast flow will have a BFS tree (down-link direction) rooted at the base station or a shortest path (up-link direction) connecting the source node to the base station in each cell if this cell contains at least one receiver of this flow. Due to the result in [9], we know that for each flow, the broadcast capacity achieved by the BFS tree constructed in Algorithm 1 is $\Theta(W_a)$ and it is not difficult to show that the up-link direction shortest path which connects the source node to the base station can achieve rate $\Theta(W_a)$ as well without considering all other non-related simultaneously transmission. In addition, by Lemma 19, we know when the total number of multicast flow n_s satisfies the condition 7, the maximum number of ad hoc flows inside of any cell satisfies $\max_{j=1}^{m} f'_j$ is $\Theta(n_s \frac{k}{m})$. Assume $\max_{j=1}^{m} f'_j = c_8 n_s \frac{k}{m}$ for some constant c_8. In other words, *w.h.p*, we have at most have $c_8 n_s \frac{k}{m}$ up-link flows or $c_8 n_s \frac{k}{m}$ down-link flows existing in each cell. We simply consider the down-link flows and up-link flows separately. Clearly, by a TDMA schedule, the minimum per-flow rate for both up-link flows and down-link flows can reach at least $\frac{W_a}{c_8 n_s \frac{k}{m}}$. Hence, we have

Theorem 24: When $k = O(m)$ and n_s satisfies the condition 7, the lower bound capacity for ad hoc links achieved by applying Algorithm 1 and TDMA schedule is $\Omega(\frac{W_a m}{n_s k})$ With probability at least $1-\frac{1}{n}$.

C. When $k = O(a^2/r^2)$ and $k = \Omega(m)$

We still use Algorithm 1 to do routing. The achievable lower bound capacity for backbone links and cellular links are easy to get (similar analysis as we did when $k = O(m)$). The only difference in this case is that the number of multicast flows which will go through some base station could be up to but no more than n_s flows due to Lemma 21. Then after applying Algorithm 1 and TDMA scheduling, the achievable lower bound capacity by backbone links and cellular links are $\Omega(\frac{W_B}{n_s})$ and $\Omega(\frac{W_c}{n_s})$ respectively.

The lower bound capacity achievable by all ad hoc links using *Cellular Routing* (Algorithm 1) when $k = \Omega(m)$ can be get by the similar proof as we did in subsection IV-B. The difference is that the possible up-link flows and down-link flows in each cell could be up to but no more than n_s flows (due to Lemma 21). By the same argument, we have the following theorem.

Theorem 25: With probability at least $1-\frac{1}{n}$, the minimum per-flow rate for ad hoc links achievable by applying Algorithm 1 is $\Omega(\frac{W_a}{n_s})$ when $k = \Omega(m)$ and $k = O(\frac{a^2}{r^2})$.

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

In this case, we have proved that the upper bound on the total multicast capacity is only $\Theta(W)$. Obviously, the total multicast capacity for hybrid network is at least the lower bound of the capacity for broadcast no matter we use either *Cellular Routing* or *Ad Hoc Routing*. In [9], Keshavarz-Haddad *et al.*. present a broadcast scheme to achieve capacity $\Theta(W_a)$. Thus, we have

Theorem 26: The minimum per-flow multicast capacity achievable by all ad hoc links is at least $c_7\frac{W_a}{n_s}$, where $c_7 \leq \frac{1}{\Delta+1}$ is a constant.

Obviously, the minimum per-flow multicast capacity achievable by backbone links and cellular links are $\Omega(\frac{W_B}{n_s})$ and $\Omega(\frac{W_c}{n_s})$ by the similar analysis we used in IV-C.

V. Capacity Bound for Hybrid Routing

In this section, we will give asymptotic upper bounds for any *Hybrid Routing* strategy. The surprising implication of this results is that if we choose the one who can gain larger capacity between *Ad Hoc Routing* and *Cellular Routing* as our routing strategy, the attainable capacity is the same order of the upper bounds of any *Hybrid Routing* strategy. It implies that the upper bounds are tight and our routing strategy is asymptotically optimal.

The aforementioned result is based on the following observation. First, when *Hybrid Routing* is applied, any link in G could be used, in other words, for any multicast flow f_i, the corresponding (resultant) multicast tree T_i for *Hybrid Routing* may contain at most three types of links, the links in E_a, E_B or E_c. Assume $T_a^i = E_a \bigcap T_i$, $T_B^i = E_B \bigcap T_i$ and $T_c^i = E_c \bigcap T_i$, *i.e.*, T_a^i (T_B^i and T_c^i) contains all ad hoc links (backbone links and cellular links) used by tree T_i. Furthermore, we use sets T_a, T_B and T_c to denote the "union" of all ad hoc links, backbone links and cellular links used by all n_s multicast trees. Notice, here the reason that we quote the word union is because the link which is belong to different multicast trees will be counted multiple times in T_a, T_B and T_c, *i.e.*, if we use $|S|$ to denote the summation of the ' length of all links belong to link set S, then $|T_a| = \sum_{i=1}^{n_s}|T_a^i|$, $|T_B| = \sum_{i=1}^{n_s}|T_B^i|$, $|T_c| = \sum_{i=1}^{n_s}|T_c^i|$.

A. When $k = O(m)$

Instead of studying the upper bound for any giving routing strategy directly, we may view this problem in a alternative way: For any given routing strategy, if we can always construct a new routing tree based on it such that the upper bound of multicast capacity by using our new routing strategy is no smaller than the original one, then the upper bounds for the new routing strategy must be one of the valid upper bounds for the original routing strategy. Next, we first give a illustration of our construction method, then a upper bound for the new constructed routing tree will be derived, finally, we use the above upper bounds as one desired upper bounds. In the following contents, we will use D_i to denote a set of base stations used by flow i such that each base station in this set has at least one cellular link adjacent to it.

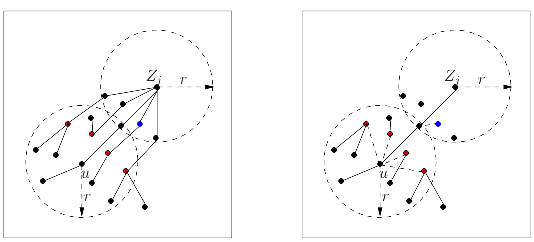
We first give a illustration of our construction approach based on the given routing strategy T_i for flow i as follows:

1) Use the minimum length tree spanning D_i to replace T_B^i, we use $T_B^{i\prime}$ to denote new tree.
2) Adjust the links contained in T_c^i and T_a^i such that there are no more than 19 cellular links on the resultant tree after construction to each base station, denoting the new trees (forests) by $T_c^{i\prime}$ and $T_a^{i\prime}$ respectively .

Next we will explain and analyze these stages in details, in the following contents, we will use λ_a^i, $\lambda_a^{i\prime}$, λ_B^i, $\lambda_B^{i\prime}$, λ_c^i and $\lambda_c^{i\prime}$ to denote the achievable data rate on T_a^i, $T_a^{i\prime}$, T_B^i, $T_B^{i\prime}$, T_c^i and $T_c^{i\prime}$ respectively.

First, we have $|T_B^{i\prime}| \leq |T_B^i|$, it is straight forward from the fact that $T_B^{i\prime}$ is a minimum length tree spanning D_i. Second, we have $\lambda_c^{i\prime} \geq c_{11}\lambda_c^i$ for some constant c_{11}, this is based on the following observation: Due to the results in [23], we know we can find at most 19 nodes as "connectors"(one hop away from the base station using cellular links) which can connect to a number of wireless nodes (say "dominators", two hop away from the base station) such that these dominators can cover all nodes which are two hop away from base station. Next, we let all receivers (exiting in the communication disk of the base station) which are not selected to the 19 base stations connect to the closest connector. On the one hand, for down-link from the base station, our construction will not decrease the cellular link rate. On the other hand, for upper link, we know that only constant number of wireless node (in our case at most 4κ nodes) can transmit at the same time such that each of 19 connectors will have addition burden at most 4κ times than before the construction. Obviously, comparing with the original scheduling period $\mathcal{T}$, after construction, $4\kappa\mathcal{T}$ time is enough for a scheduling period. Hence, we have $\lambda_c^{i\prime} \geq c_{11}\lambda_c^i$ for some constant c_{11}. See Fig. 4 for illustration. Third, $\lambda_a^{i\prime} \geq c_{10}\lambda_a^i$ for some constant c_{10}. We guarantee this point by the following reservation. After we get connectors during the second step (after adjust cellular links). For some nodes which are two hop away from the base station in the routing tree before our construction, they could lost the connection to the base station when their relaying nodes to the base station are not selected as "connectors" in the second step. If so, we simply let these nodes connect to the closest "dominator". See Fig. 4 for illustration. Let us take u as an example. First, for each internal node u ("dominator") closed to the base station. After reconstruction, some other nodes (who lost connection to the base station) will turn to u and ask u to help to relay traffic to the base station. However, these nodes must satisfy two conditions. (1). They are in the communication range of u. (2). They can transmit simultaneously based on the original routing and scheduling strategy. Clearly, u can at most cover 4 closest square region with side length r and for each square cell with side length r, there are at most κ nodes can transmit simultaneously as we have proved before. Hence, we can guarantee, there are at most 4κ nodes will turn to u in one time slot after construction, in other words, after construction, node u can achieve at least $\frac{1}{4\kappa}$ rate of the original rate before construction by a TDMA scheduling, *i.e.*, ${\lambda_a^i}' \geq \frac{1}{4\kappa}\lambda_a^i$.

These three important observations guarantee that the upper bounds for the new constructed routing strategy derived from the following analysis are also valid upper bounds for the original routing strategy. See Fig. 4 for illustration.



(a) Before reconstruction. (b) After reconstruction.

Fig. 4. Part of the Hybrid Network which is near the base station. The dash line from red nodes to u are new added "burden" for node u. z_j is the base station for j^{th} cell. Black nodes denote internal nodes. Red nodes are nodes who will changed their routing strategy after construction. Blue node is one of receivers which is in the communication disk of the base station z_j.

From now on, we focus on studying the upper bounds for the new constructed routing tree: Due to the result in [11], we know that the total length of internal edges of n_s multicast trees spanning k receivers satisfies

$\sum_{i=1}^{n_s}|T_i'| = \sum_{i=1}^{n_s}|T_a^{i'}| + \sum_{i=1}^{n_s}|{T_B^i}'| + \sum_{i=1}^{n_s}|T_c^{i'}| = |T_a'| + |T_B'| + |T_c'| \geq c_9 n_s \cdot a \cdot \sqrt{k}$

for some constant c_9. Then we discuss the following two cases respectively:

(1) If $|T_a'| \geq |T_B'| + |T_c'|$: Since the total length of T_a', T_B' and T_c' is no smaller than $c_9 n_s \cdot a \cdot \sqrt{k}$, we have $|T_a'| \geq \frac{1}{2}c_9 n_s \cdot a \cdot \sqrt{k}$. As shown in [11], the total area covered by all of these ad hoc trees is at least $\eta_1 \cdot n_s \cdot a \cdot \sqrt{k} \cdot r$ for some constant η_1, the number of nodes covered by all ad hoc trees is at least $\eta_1 n_s \cdot a\sqrt{k}r \times \frac{n}{a^2}$ with high probability. Based on the data copy argument, it follows that: $\lambda_a^i \leq c_{10}{\lambda_a^i}' \leq \frac{c_{10}W_a a}{\eta_1 n_s\sqrt{k}r}$.

(2) If $|T_a'| < |T_B'| + |T_c'|$: We have $|T_B'| + |T_c'| \geq \frac{1}{2}c_9 n_s \cdot a \cdot \sqrt{k}$ According our construction approach, we know that for any routing tree T_i', each base station has at most 19 adjacent cellular links, then together with the fact that the length of each cellular link is at most r, we have the following inequality: $|T_c'| \leq \frac{|T_B'|}{a/\sqrt{m}} \times 19 \times r$. Since $m = O(\frac{a^2}{r^2})$, we have $|T_c'| \leq |T_B'|/19$, it follows that $|T_B'| \geq c_3 n_s a\sqrt{k}$ for some constant c_3. This implies that there is at least one base station which is used by at least $\frac{c_3 n_s a\sqrt{a}}{\sqrt{m}}$ flows. Due to the congestion on this base station, we gain the following upper bound: $\lambda_B^i \leq {\lambda_B^i}' \leq \frac{W_B\sqrt{m}}{c_3 n_s\sqrt{k}}$. Furthermore, because ${T_B^i}'$ spans $|D_i|$ base stations, we have $|{T_B^i}'| \leq 2\sqrt{2}\sqrt{|D_i|}a$ according to the results in [11]. It follows that $c_3 n_s a\sqrt{k} \leq |T_B'| = \sum_{i=1}^{n_s}|{T_B^i}'| \leq \sum_{i=1}^{n_s} 2\sqrt{2}\sqrt{|D_i|}a$.

Together with the fact that $(\sum a_i^p)^{\frac{1}{p}}(\sum b_i^q)^{\frac{1}{q}} \geq \sum a_i b_i$, we have the following inequality:

$$\sum_{i=1}^{n_s}|D_i| \geq \eta_2(\frac{\sum_{i=1}^{n_s}\sqrt{|D_i|}}{\sqrt{n_s}})^2 \geq (\frac{\eta_3 n_s\sqrt{k}}{\sqrt{n_s}})^2 = \eta_3^2 n_s k$$

We conclude that there is at least one based station which is used by at least $\frac{\eta_3^2 n_s k}{m}$ flows to connect wireless nodes directly. Due to the congestion on both ad hoc links and cellular links accessing the base station, we further have the following two upper bounds for this case:

$$\begin{cases} \lambda_a^i \leq c_{10}{\lambda_a^i}' \leq \frac{W_a}{\eta_3^2 n_s k/m} = \frac{W_a m}{\eta_3^2 n_s k} \\ \lambda_c^i \leq c_{11}{\lambda_c^i}' \leq \frac{W_c}{\eta_3^2 n_s k/m} = \frac{W_c m}{\eta_3^2 n_s k} \end{cases} \quad (9)$$

It concludes that the upper bound in this case is

$$O(\min\{\frac{W_B\sqrt{m}}{n_s\sqrt{k}}, \frac{W_a m}{n_s k}, \frac{W_b m}{n_s k}\})$$

The final upper bound is gained by choosing the maximum one between case 1) and case 2):

Lemma 27: The capacity bound for any *Hybrid Routing* strategy is

$$O(\max\left[\min\{\frac{W_B\sqrt{m}}{n_s\sqrt{k}}, \frac{W_a m}{n_s k}, \frac{W_b m}{n_s k}\}, \frac{W_a a}{n_s\sqrt{k}r})\right])$$

when $k = \Omega(m)$ and $m = O(\frac{a^2}{r^2})$.

It implies that when $k = \Omega(m)$, its asymptotic optimal to choose the larger one between *Ad Hoc Routing* and *Cellular Routing* as our routing strategy based on the calculated lower bound for each routing strategy.

B. When $k = \Omega(m)$ and $k = O(\frac{a^2}{r^2})$

Same as the proof for the previous case, we first construct a new routing tree based on any given routing strategy. Since the upper bound for the new constructed routing tree can also be considered as a valid upper bound for the original routing strategy, we will focus on studying the upper bound for the new constructed routing tree.

Similarly, we have two possible cases need to address:

(1) If $|T_a'| \geq |T_B'| + |T_c'|$: The proof is exactly same as the one shown before for the same case, we gain following upper bound: $\lambda_a^i \leq c_{10}{\lambda_a^i}' \leq \frac{c_{10}W_a a}{\eta_1 n_s\sqrt{k}r}$.

(2) If $|T'_a| < |T'_B| + |T'_c|$: We will prove that this case is impossible when $k = \Omega(m)$. Since ${T^i_B}'$ is a tree spanning at most m base stations using only backbone links, we get $|T'_B| \leq n_s(m-1) \cdot a/\sqrt{m} < n_s\sqrt{m}a$, we also know that $|T'_c| \leq 19n_s mr$ because each base station has at most 19 adjacent cellular links. We immediately have when $k > \Omega(m)$, it is impossible that $|T'_B| + |T'_c| > \frac{1}{2}c_9 n_s \cdot a \cdot \sqrt{k}$, in other words, $|T'_a| \not< |T'_B| + |T'_c|$.

We finally have the following lemma:

Lemma 28: The capacity bound for any *Hybrid Routing* strategy is $O(\frac{W_a a}{n_s\sqrt{k}r})$ when $k = \Omega(m)$, $k = O(\frac{a^2}{r^2})$ and $m = O(\frac{a^2}{r^2})$.

This result implies that when $k = \Omega(m)$ and $k = O(\frac{a^2}{r^2})$, using *Ad Hoc Routing* is already asymptotic optimum.

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

Again, because when $k = \Omega(\frac{a^2}{r^2})$, the union of the transmission disks of these k receiver nodes in a multicast session will cover at least a constant fraction, say $0 < \rho_2 \leq 1$, of the deployment region with high probability when $k = \Omega(\frac{a^2}{r^2})$. Then based on the data copy argument stated in [12], we have the following lemma:

Lemma 29: The capacity bounds for any *Hybrid Routing* strategy is $O(\frac{W_a}{n_s})$ when $k = \Omega(\frac{a^2}{r^2})$, $m = O(\frac{a^2}{r^2})$.

It is not hard to find that its asymptotic optimum to choose *Ad Hoc Routing* as our routing strategy when $k = O(\frac{a^2}{r^2})$.

D. Put It All Together

By summarizing these results, we have

Theorem 30: The upper bounds of the multicast capacity for any *Hybrid Routing* strategy is

$$\begin{cases} O(\max\left[\min\left(\frac{W_B\sqrt{m}}{n_s\sqrt{k}}, \frac{W_b m}{n_s k}, \frac{W_a m}{n_s k}\right), \frac{W_a}{n_s\sqrt{k}}\frac{a}{r}\right]) & \text{if } k = O(m) \\ O(\frac{a}{r}\cdot\frac{W_a}{n_s\sqrt{k}}) \quad \text{if } k = \Omega(m), k = O(\frac{a^2}{r^2}) \\ O(\frac{W_a}{n_s}) \quad \text{if } k = \Omega(\frac{a^2}{r^2}) \end{cases} \tag{10}$$

We then give a general routing strategy that can achieve the asymptotic upper bound for hybrid network $N_{n,m,a}$:

- If $k = O(m)$, we choose the one who can gain larger data rate between *Ad Hoc Routing* strategy and *Cellular Routing* strategy.
- If $k = \Omega(m)$, we use *Ad Hoc Routing* strategy.

Then together with the lower bound for *Ad Hoc Routing* strategy and *Cellular Routing* strategy, we get the main theorem 2.

VI. Literature Review

Gupta and Kumar [6] studied the asymptotic unicast capacity of a multi-hop wireless networks. When each wireless node is capable of transmitting at W-bps using a constant transmission range, the throughput achievable by *each* node for a randomly chosen destination is $\Theta(\frac{W}{\sqrt{n\log n}})$ bits per second. Grossglauser and Tse recently showed that the unicast capacity can be improved by the mobility of wireless nodes regardless of delay. Gastpar and Vetterli studied the capacity of random networks using relay in [1]. Chuah *et al.* [2] studied the capacity scaling in MIMO wireless systems under correlated fading. The capacity scaling in delay tolerant networks with heterogeneous mobile devices was studied by Garetto *et al.* [3]. Keshavarz-Haddad *et al.* studied the bounds for the capacity of wireless networks imposed by topology and demand in [8]. Their techniques can be used to study unicast, multicast and broadcast capacity.

Broadcast capacity of both arbitrary networks and random networks has been studied in [9, 14]. Keshavarz-Haddad *et al.* [4] studied the broadcast capacity with dynamic power adjustment for physical interference model.

Multicast capacity was also studied in the literature. Jacquet and Rodolakis [5] 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 $O(\frac{W}{\sqrt{kn\log n}})$. Recently, rigorous proofs of the multicast capacity were given in [12, 15]. Li *et al.* [12] studied asymptotic multicast capacity for a large-scale random wireless networks. 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.*, $\Theta(W)$. Li *et al.* [18] studied the lower bound of multicast capacity for large scale wireless networks under Gaussian Channel model by presenting some novel methods. This result was recently improved by Wang *et al.* [24]. Hu *et al.* [10] recently studied the capacity and delay tradeoffs of multicast capacity when the mobility model is i.i.d. They show that mobility and redundancy do improve the multicast capacity when the number of receivers k per flow is small. Lee *et al.* [20] studied the scalability of DTN multicast routing. They propose RelayCast, a routing scheme that extends the two-hop relay algorithm in [7] in the multicast scenario.

Liu *et al.* [17] studied the unicast capacity of hybrid network (a wireless ad hoc network with infrastructure). They essentially studied the unicast capacity of hybrid wireless networks under the one-dimensional network model and two-dimensional strip model respectively. Kozat and Tassiulas [16] also studied the unicast capacity of ad hoc networks with a random flat topology under the present support of an infinite capacity infrastructure network. They showed that the per source node capacity of $\Theta(W/\log n)$. In [19], Mao *et al.*, studied the multicast capacity for hybrid networks by using *Cellular Routing* strategy.

VII. Conclusions

In this paper, we essentially studied the multicast capacity that can be achieved by hybrid networks with randomly distributed wireless nodes and regularly distributed base stations. We derived analytical upper bounds and lower bounds on multicast capacity of hybrid 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 fixed square with side length $a = 1$, while the transmission range is selected appropriately, *i.e.*, $r = \Theta(\sqrt{\frac{\log n}{c\cdot n}})$ for some constant c. In addition, our results still hold when $r = 1$ while the deployment region has a bounded aspect ratio such as a disk or a rectangular area when ratio width/height is bounded.

A number of interesting questions remain challenging. The first is to study the capacity when other interference models are applied such as physical interference model and Gaussian channel model. The second is to investigate the capacity regions when opportunistic spectrum usage is adopted by some wireless terminals. The last, but not the least, is to study the capacity region for delay tolerant networks where the wireless terminals are mobile (following some mobility model) and we want to study the achievable capacity when we can tolerate certain delay.

REFERENCES

[1] GASTPAR, M., AND VETTERLI, M. On the capacity of wireless networks: the relay case. In *IEEE INFOCOM* (2002).

[2] CHUAH, C.-N., TSE, D. N. C., KAHN, J. M., AND VALENZUELA, R. A. Capacity scaling in mimo wireless systems under correlated fading. In *IEEE Transactions On Information Theory* (2002), vol. 48.

[3] GARETTO, M., GIACCONE, P., AND LEONARDI, E. Capacity scaling in delay tolerant networks with heterogeneous mobile nodes. In *ACM MobiHoc* (2007), pp. 41–50.

[4] KESHAVARZ-HADDAD, A., AND RIEDI, R. On the broadcast capacity of multihop wireless networks: Interplay of power, density and interference. In *IEEE (SECON)* (2007).

[5] JACQUET, P., AND RODOLAKIS, G. Multicast scaling properties in massively dense ad hoc networks. In *ICPADS* (2005), pp. 93–99.

[6] GUPTA, P., AND KUMAR, P. Capacity of wireless networks. *IEEE Transactions on Information Theory IT-46* (1999), 388–404.

[7] GROSSGLAUSER, M., AND TSE, D. Mobility increases the capacity of ad hoc wireless networks. *IEEE/ACM TON 10*, 4 (2002), 477–486.

[8] KESHAVARZ-HADDAD, A., AND RIEDI, R. H. Bounds for the capacity of wireless multihop networks imposed by topology and demand. In *ACM MobiHoc* (2007), pp. 256–265.

[9] KESHAVARZ-HADDAD, A., RIBEIRO, V., AND RIEDI, R. Broadcast capacity in multihop wireless networks. In *ACM MobiCom* (2006), pp. 239–250.

[10] CHENHUI HU, XINBING WANG, AND FENG WU MotionCast: On the Capacity and Delay Tradeoffs ACM Mobihoc 2009

[11] LI, X.-Y., Multicast Capacity of Wireless Ad Hoc Networks, In *IEEE/ACM Transaction on Networking* (2008), to appear.

[12] LI, X.-Y., TANG, S.-J., AND OPHIR, F. Multicast capacity for 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] ULA C. KOZAT AND LEANDROS TASSIULAS, Throughput capacity of random ad hoc networks with infrastructure support, *Proc. ACM Mobihoc*, 2003.

[17] BENYUAN LIU, PATRICK THIRAN, DON TOWSLEY, Capacity of a Wireless Ad Hoc Network with Infrastructure, *ACM Mobihoc*, 2007.

[18] SHI LI, YUNHAO LIU, XIANG-YANG LI, Capacity of Large Scale Wireless Networks Under Gaussian Channel Model *Proc. ACM MobiCom*, 2008.

[19] XUFEI MAO, XIANG-YANG LI, AND SHAOJIE TANG, Multicast Capacity for Hybrid Wireless networks *Proc. ACM MobiHoc*, 2008.

[20] UICHIN LEE, S.Y. OH, K.-W. LEE, AND M. GERLA RelayCast: Scalable Multicast Routing in Delay Tolerant Networks, IEEE ICNP 2008.

[21] ABU-MOSTAFA,YASER S, The Vapnik-Chervonenkis Dimension: Information versus Complexity in Learning *Neural Computation*, 1989.

[22] J. M. STEELE, Growth rates of euclidean minimal spanning trees with power weighted edges. *The Annals of Probability, 16(4):17671787* , 1988

[23] SCOTT C.-H. HUANG, PENG-JUN WAN, CHINH T. VU, YINGSHU LI, F. FRANCES YAO, Nearly Constant Approximation for Data Aggregation Scheduling in Wireless Sensor Networks. *INFOCOM 366-372* , 2007

[24] CHEN WANG, XIANG-YANG LI, CHANGJUN JIANG, SHAO-JIE TANG, YUNHAO LIU, AND JIZHONG ZHAO Scaling laws of networking-theoretic bounds on capacity for wireless networks, IEEE INFOCOM 2009