Capacity of a Wireless Ad Hoc Network with Infrastructure

Benyuan Liu [1], Zhen Liu [2], Don Towsley [3]

[1] Department of Computer Science

University of Massachusetts Lowell

Lowell, MA 01854

[2] IBM T.J. Watson Research Center, P.O. Box 704

Yorktown Heights, NY 10598

[3] Department of Computer Science

University of Massachusetts Amherst

Amherst, MA 01002

Abstract

We study the capacity of a wireless ad hoc network with infrastructure, where well-connected base stations are placed in an ad hoc network to relay data traffic for wireless nodes. This network architecture presents a tradeoff between a cellular network and an ad hoc network in that data may be forwarded in a multi-hop fashion or through the infrastructure. It has been shown that the capacity of ad hoc network does not scale well with the number of nodes in the system [1]. In this work, we investigate the potential benefit of infrastructure network to improve ad hoc network capacity. Analytical expressions are obtained for the capacity of an ad hoc networks with infrastructure. For an ad hoc network of n nodes with m base stations, the results show that, if m grows asymptotically slower than $\sqrt{n}$, the benefit of infrastructure network is insignificant. However, if m grows faster than $\sqrt{n}$, the capacity increases linearly with the number of base stations, providing an effective improvement over ad hoc networks. Therefore, in order to achieve non-negligible capacity gain, the investment in the wired infrastructure should be sufficiently high.

I. INTRODUCTION

Network capacity is a key characteristic of a wireless network. It represents the long-term achievable data transmission rate that a network can support. The capacity of a wireless network depends on many aspects of the network: network architecture, power and bandwidth constraints, communication paradigm, routing strategy, and radio interference model. A good understanding of the capacities of different network architectures allows a designer to choose an architecture appropriate for his or her specific purpose.

Several network architectures have been proposed for wireless data networks. In a wireless cellular network or a wireless LAN, nodes communicate with each other through base stations or access points. A node first connects to the nearest base station or access point in order to communicate with other nodes. In this scenario, base stations (access points) serve as communication gateways for all the nodes in its cell (basic service area).

In situations where there is no fixed infrastructure, for example, battle fields, catastrophe control, etc, wireless ad hoc networks become viable alternatives to wireless cellular networks or wireless LANs for nodes to communicate with each other. A wireless ad hoc network is a communication network formed by a collection of nodes without the aid of any fixed infrastructure. In a wireless ad hoc network, due to the lack of infrastructure and the limited transmission range of each node, data needs to be routed from the source to the destination by peer nodes in a multi-hop fashion.

Recently, a hybrid of wireless ad hoc and cellular network architectures has been proposed to improve various aspects of the performance of wireless ad hoc or cellular networks, for example, cell coverage, system throughput, power efficiency, network connectivity and reliability [2], [3], [4], [5], [6]. The hybrid network architecture consists of normal wireless nodes and a number of well-connected base stations, which are regularly or randomly placed in the network. The network is divided into cells, each containing a base station. Figure 1 depicts the scenario where base stations are regularly placed in a hexagon pattern. The base stations neither produce data nor consume data. They act solely as relay nodes to improve network performance and only engage in routing and forwarding data for normal wireless nodes. In this work, we assume all nodes and base stations are static; no node and base station can move. In a wireless ad hoc network with infrastructure (well-connected base stations), data may be forwarded in a multi-hop fashion using the wireless node as in a wireless ad hoc network, or through the infrastructure as in a cellular network. Communications using multi-hop forwarding and communications using
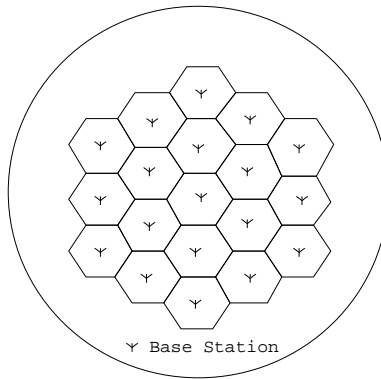
Fig. 1.   Wireless Ad hoc Network with Infrastructure

the infrastructure coexist in the network. Although some system aspects (throughput, routing, reliability) have been studied for the hybrid network architecture, little is known in terms of its network capacity. It is of great interest to find out what capacity gains can be achieved by adding infrastructure to a wireless ad hoc network.

In [1], Gupta and Kumar show that the per-node capacity of a random ad hoc network scales as $\Theta(\frac{W}{\sqrt{n\log n}})$, where n is the number of nodes in the network. In the limit, the per-node capacity diminishes to zero as the number of node, n, goes to infinity. However, it is not clear how much capacity gain can be achieved by adding infrastructure to a wireless ad hoc network. Intuitively, on one hand, data can be partly routed through the infrastructure, which helps to reduce the number of wireless transmissions. The network can accommodate more data and, hence, a higher capacity can be achieved. On the other hand, too much use of the infrastructure may cause hot spots around base stations and poor spatial concurrency, leading to a reduced capacity. The goal of this work is to study the capacity of a wireless ad hoc network with infrastructure. In particular, we are interested in the following two questions:

- How does the capacity of a wireless ad hoc network with infrastructure scale with the number of nodes and the number of base stations?
- What capacity gain can be achieved by adding infrastructure to a wireless ad hoc network?

In this work, we assume a wireless ad hoc network of n nodes with an infrastructure network of m base stations. Each node and base station is capable of transmitting at a rate of W bits/sec over the common wireless channel. We consider a deterministic routing scheme. A node sends data through the infrastructure if the destination is outside of the cell where the source is located.

Otherwise, the data are forwarded in a multi-hop fashion without using infrastructure as in an ad hoc network. Intuitively, the deterministic routing strategy improves network capacity in the following two ways. First, allowing sessions whose end points do not fall within the same cell to use the infrastructure reduces the number of wireless transmissions. The network can accommodate more data and hence the capacity increases. Second, requiring sessions whose end points lie within the same cell to communicate in an ad hoc mode utilizes the available spatial concurrency, which further improves network capacity.

We show that if base stations are regularly placed, under the deterministic routing strategy, if m, the number of base stations, grows slower than $\sqrt{n}$, the capacity scales as $\Theta\left(\sqrt{\frac{n}{\log(n/m^2)}}\right)$. In this case, compared to the capacity of a pure wireless ad hoc network without infrastructure, the benefit of adding base stations is insignificant. However, if base stations can be added at a rate faster than $\sqrt{n}$, the capacity scales as $\Theta(m)$. The capacity increases linearly as more base stations are added, presenting an effective improvement over the original wireless ad hoc network. Therefore, in order to achieve a non-negligible capacity gain, the investment in the wired base station infrastructure should be high enough: the number of base stations should grow at a rate faster than $\sqrt{n}$, as the number of wireless nodes increases. Similar results are also obtained for the case where base stations are randomly deployed.

In establishing the capacity of a wireless ad hoc network with infrastructure, we also derive the capacity of an inhomogeneous wireless ad hoc network, where not every node has data to send, and not every node can be a receiver. This is a generalization of the results of Gupta and Kumar for a homogeneous wireless network, where every node has data to send to a randomly chosen location, and every node is potentially a receiver. The capacity of an inhomogeneous wireless ad hoc network is of great interest in its own right, since it is present in many application scenarios. For example, in some sensor network applications, only those sensors that sense interesting events will have data to send, and data sinks are usually chosen from a set of designated nodes in the network.

The rest of the paper is organized as follows. In Section II, we review previous results on the capacity of wireless ad hoc networks. In Section III, we describe the network model and define the capacity of the network model. In Section IV, we present the capacity of an inhomogeneous wireless ad hoc network. In Section V, we derive expressions for network capacity under the deterministic routing strategy. The proofs of theorems are presented in Section . In Section VI, we present the conclusions.

II. RELATED WORK

Recently, there has been a flurry of studies on the capacity of wireless ad hoc networks. In [1], Gupta and Kumar study the throughput capacity of a two-dimensional random wireless network, where fixed nodes are randomly placed in the network and each node sends data to a randomly chosen destination. The throughput capacity per node scales as $\Theta(\frac{1}{\sqrt{n \log n}})$, where n is the number of nodes in the network. In the limit, the per-node throughput vanishes as the number of node n goes to infinity. The capacity of a three-dimensional wireless ad hoc network has been obtained by the same authors [7]. In this case, per-node throughput scales as $\Theta\left(\frac{1}{\left(n \log^2 n\right)^{\frac{1}{3}}}\right)$. Although three-dimensional wireless networks have higher capacity than two-dimensional networks, the throughput obtained by each node still tends to zero as the number of nodes in the network increases.

In [8], Grossglauser and Tse show that the capacity of wireless ad hoc networks can be significantly improved by exploiting node mobility. Using a one-hop relaying mechanism, the proposed scheme achieves an aggregate throughput capacity of $O(n)$, or a per-node throughput capacity of $O(1)$. However, this capacity improvement is achieved at the cost of unbounded packet delay and unbounded buffer requirement at each node. To address the concern of unbounded delay, the authors of [9] propose a routing algorithm which exploits the patterns in the mobility of nodes to provide guarantees on the packet delay. They show that packet delay is small and bounded, while the throughput achieved by the algorithm is only a poly-logarithmic factor from the optimal.

The capacity of wireless ad hoc networks exhibits different behaviors for different communication paradigms. In [10], Gastpar and Vetterli study the scenario where there is only one active source and destination pair, while the other $(n-2)$ nodes serve as relays. The authors approach the problem from an information theoretical perspective and allow arbitrary network coding. In this case, the throughput capacity between the source and destination is shown to scale as $O(\log n)$, which grows logarithmically with the number of relaying nodes. In [11], the authors study a many-to-one communication paradigm where a total of n nodes send data to a single central receiver. Under a transmission and interference model similar to that of Gupta and Kumar [1], the authors show that the network capacity is $\Theta(1)$, with a per sender capacity of $\Theta(1/n)$. In [12], Li et al. examine the effect of IEEE 802.11 on network capacity. They also derive the criteria for traffic patterns that makes the capacity scale with the network size.

The benefit of using directional antennas to improve wireless network capacity has been

investigated in [13] and [14]. The main finding of [13] is that one can only achieve a factor of $\Theta\left(\log^2 n\right)$ increase in capacity by allowing arbitrarily complex signal processing at transmitters and receivers. In [14], the authors derive the capacity gain in terms of the beam-widths of transmitting and receiving antennas.

In a recent work [15], the authors study the capacity of a network architecture similar to ours. Under the assumption that the number of wireless nodes per base station is bounded above by a constant and wireless nodes are asymptotically connected without infrastructure, the authors show that the per-node throughput capacity scales as $\Theta\left(1/\log n\right)$. They also show that even when a less stringent requirement is imposed on topology connectivity, a per-node capacity arbitrarily close to $\Theta(1)$ can not be achieved. However, [15] only considers a special case where the number of base stations is on the same order as the number of normal wireless nodes. This implies that the investment in the infrastructure has to increase linearly as the number of nodes increases, which may not be practical in scenarios where the number of nodes is large and infrastructure resources are limited. In our work, we consider the general case and treat the number of base stations as a parameter. We find that the number of base stations plays a critical role in determining the network capacity. Depending on the relative number of base stations compared to the number of nodes, the network capacity exhibits different behaviors. We will compare and comment on the results of both work later in this paper.

III. NETWORK MODEL

In this section, we first describe the network architecture, interference model, data transmission modes, and routing strategies in a wireless ad hoc network with infrastructure. We then formally define the capacity of the network model.

A. Network Architecture

As in [1], we scale space and assume that a population of n nodes is randomly, independently and uniformly, located within a unit area disk in the plane. We further assume that the nodes are homogeneous, employing the same transmission range or power. Every node is a data source. The destination for each node is independently chosen as the node nearest to a randomly located point within the unit area disk. In addition to the n wireless nodes in the network, a network of m base stations is placed in the network. We consider two base station deployment strategies. In the first strategy, base stations are regularly placed in the unit area disk, dividing the area

into a hexagon tessellation. Each hexagon is called a cell, as in cellular networks. In the second strategy, base stations are randomly deployed in the network. These base stations are assumed to be connected together by an infinite bandwidth wired network. Hence, there are no bandwidth constraints among base stations. Moreover, the base stations neither produce data nor consume data. They act solely as relay nodes to improve network performance and only engage in routing and forwarding data for normal wireless nodes. All the nodes and base stations are assumed to be static; there is no mobility involved in this work.

B. Interference Model

All the nodes and base stations share a common wireless channel. We assume a time-division multiplexing (TDMA) scheme for the data transmission over the wireless channel. Time is divided into fixed duration slots. In each time slot, certain nodes are scheduled to send data. A node cannot transmit and receive data simultaneously, and a node can only receive data from one other node at the same time.

All of the nodes, including the base stations, employ the same wireless transmission range, denoted by r. For the interference model, we adopt the Protocol Model introduced in [1]. The definition of the Protocol Model is given as follows.

A transmission from node X_i is successfully received by node X_j if the following two conditions are satisfied:

1) Node X_j is within the transmission range of node X_i, i.e.,

$$|X_i - X_j| \leq r$$

where $|X_i - X_j|$ represents the distance between node X_i and node X_j in the plane.

2) For every other node X_k that is simultaneously transmitting over the same channel,

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

This condition guarantees a guard zone around the receiving node to prevent a neighboring node from transmitting on the same channel at the same time. The radius of the guard zone is $(1 + \Delta)$ times the distance between the sender and receiver. The parameter Δ defines the size of the guard zone and we require that $\Delta > 0$.

C. Transmission modes and Routing Strategies

In a wireless ad hoc network with infrastructure, there are two transmission modes: ad hoc mode and infrastructure mode. In ad hoc mode, data are forwarded from a source to the destination in a multi-hop fashion without using the infrastructure. In infrastructure mode, data are forwarded through the infrastructure. More specifically, data are first transmitted from the source to the nearest base station over the wireless channel. The base station then forwards the data through the wired infrastructure to the base station nearest to the destination, which finally transmits the data to the destination over wireless channel. Note that in the infrastructure mode, wireless communications are used for transmissions between base stations and end nodes, while transmissions between base stations are carried by the infrastructure.

We consider a deterministic routing strategy. If the destination node is located in the same cell as the source node, data are forwarded in ad hoc mode. Otherwise, if the destination node is not located in the same cell as the source node, data are forwarded in infrastructure mode. We provide the intuition behind the routing strategy: if the destination node is close, we would like to take advantage of the spatial concurrency by using ad hoc mode without bottlenecking the base stations; however, if the destination is far away, we would like to take advantage of the free bandwidth in the infrastructure to alleviate the multi-hop burden of the long route.

We assume each node can transmit W bits/sec over the common wireless channel. We divide the wireless channel so that ad hoc mode transmissions and infrastructure mode transmissions go through different sub-channels. We further divide the sub-channel for infrastructure mode transmissions into uplink and downlink portions, according to the direction of the transmissions relative to the base station. Since intra-cell traffic, uplink traffic and downlink traffic use different sub-channels, there is no interference between the three types of traffics. The bandwidth assigned to intra-cell, uplink, and downlink sub-channels are W_1, W_2, and W_3, respectively, and we let $W_2 = W_3$. The transmission rates should sum to W, i.e., $\sum_{i=1}^{3} W_i = W$.

D. Definition of Capacity

To make the formulas more concise, we present the aggregate capacity of the whole network instead of the per-node capacity. Note that the per-node capacity is simply $1/n$ of the aggregate capacity of the whole network. Throughout this paper, sometimes we use capacity to refer to the aggregate capacity for the sake of conciseness. But we always use per-node capacity when we refer to the capacity per node to avoid ambiguity.

In this work, we adopt the asymptotic notations introduced in [16]. The formal definitions of feasible aggregate throughput and aggregate capacity of a wireless ad hoc network with infrastructure are presented as follows.

Definition 1: Feasible Aggregate Throughput. For a wireless ad hoc network of n nodes with an infrastructure network of m base stations, an aggregate throughput of $T(n, m)$ bits/sec is feasible if, by transmitting data in the ad hoc or infrastructure mode, there is a spatial and temporal scheduling scheme that yields an aggregate network throughput of $T(n, m)$ bits/sec on average. Here the aggregate throughput is the sum of the individual throughputs from each source node to its chosen destination.

Definition 2: Aggregate Capacity of a Wireless Ad hoc Network with Infrastructure. The aggregate capacity of a wireless ad hoc network with infrastructure is $\Theta(f(n, m))$ bits/sec if there are deterministic constants $c > 0$, and $c' < +\infty$ such that

$$\lim_{n \to \infty} \text{Prob}(\text{T(n,m) = cf(n,m) is feasible}) = 1$$

$$\lim_{n \to \infty} \inf \text{Prob}(\text{T(n,m) = } c'f(n, m) \text{ is feasible}) < 1$$

IV. INHOMOGENEOUS WIRELESS AD HOC NETWORKS

In the original work of Gupta and Kumar [1], the authors considered a homogeneous ad hoc network, where every node has data to send to a randomly chosen location, and every node is potentially a receiver. In this section, we consider an inhomogeneous ad hoc network, where not every node has data to send, and not every node can be a receiver. Such inhomogeneous networks are present in many application scenarios. For example, in some sensor network applications, only those sensors that sense interesting events will have data to send, and data sinks are usually chosen from some designated nodes in the network. Note that in [11], the authors studied a specific inhomogeneous network, namely, the many-to-one scenario, where all sensors send data to a single receiver. Our study of the capacity of the inhomogeneous network is motivated in order to derive capacity results for networks with infrastructure in Section V, however, we feel that the scenario is of interest in its own right and present it here.

We assume k nodes are independently and uniformly located within a unit-area disk, as in [1]. However, among the k nodes, there are s active source nodes and d destination nodes, where $1 \leq s, d \leq k$. Note that a node can be both a source node and a destination node, and there are not necessarily equal numbers of data sources and destinations. The source and destination

nodes are randomly located within the network. The s source nodes form a source node set S, and the d destination nodes form a destination node set D. For each source node $i \in S$, the corresponding destination node is randomly chosen from destination set D. There are s active sessions in the network. Denote the aggregate throughput capacity of an ad hoc network under this inhomogeneous communication paradigm as $T_a(k, s, d)$. The corresponding per-session throughput capacity is $\lambda_a(k, s, d) = T_a(k, s, d)/s$.

Theorem 1: For a random ad hoc network of k nodes, among which there are s unique data sources and d unique data sinks, the aggregate throughput capacity of the network is

$$T_a(k, s, d) = \Theta \left[\min \left(\min(s, d)W, \sqrt{\frac{k}{\log k}}W \right) \right] \tag{1}$$

Due to the page limit, the proof of the above theorem is presented in a detailed technical report version of this paper [17]. From Equation (1), we observe that the capacity exhibits different behaviors according to the relative number of source and destination nodes to the total number of wireless nodes. In the following, we discuss the capacity behaviors for two different regimes of number of active sources and destinations. We will see that the capacity for the two regimes are limited by different factors, thus exhibiting different behaviors. For convenience, we assume there are equal numbers of source and destination nodes, i.e., $s = d$.

- $s = O\left(\sqrt{\frac{k}{\log k}}\right)$ The network capacity is $\Theta(sW)$, and the per session capacity is $\Theta(W)$. In this case, the spatial concurrency of the network is not fully utilized, which means there are regions in the network where no transmission is taking place. In the proof, we see that there exists a spatial scheduling policy such that each source node gets one slot to transmit data in every constant number of slots. Hence each source node can send data at a constant rate. The network capacity is limited by the number of source and destination nodes.

- $s = \omega\left(\sqrt{\frac{k}{\log k}}\right)$ The network capacity is $\Theta\left(\sqrt{\frac{k}{\log k}}W\right)$, and the per session capacity is $\Theta\left(\sqrt{\frac{k}{s^2 \log k}}W\right)$. In this case, the spatial concurrency is fully utilized, and there can be no more simultaneous transmissions in the network. The capacity is limited by the maximum amount of traffic that can be supported by the network, $\sqrt{\frac{k}{\log k}}W$, as derived in [1].

V. Capacity of Wireless Ad hoc Networks with Infrastructure

In this section, we first investigate the effect of interference between wireless transmissions to the network capacity. We then derive the capacity of a wireless ad hoc network with infrastructure under the deterministic routing strategy. We also study how much capacity gain can be achieved by adding infrastructure to a wireless ad hoc network.
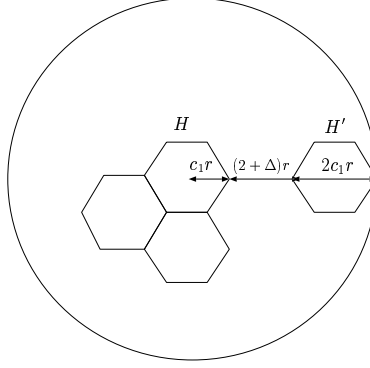
Fig. 2.   Interfering neighbors of a cell H

A. Effect of Interference

Since the intra-cell, uplink and downlink traffics are transmitted in three different sub-channels, there is no interference between the three types of traffic. However, within a sub-channel, interference exists between the same type of traffic in neighboring cells. Fortunately, the effect of this interference is minimal. We will prove that there exists a spatial transmission schedule which yields efficient frequency reuse. More specifically, the cells can be spatially divided into a constant number of different groups. Transmissions in the cells of the same group do not interfere with each other. If the groups are scheduled to transmit in a round robin fashion, each cell will be able to periodically transmit without interfering with each other. The degradation of network capacity due to the interference between the same types of traffic is thus bounded by a constant factor. We now present the formal statements and proofs.

We adopt the notion of interfering neighbors introduced in [1], and compute the number of cells that can be affected by a transmission in one cell. Two cells are defined to be interfering neighbors if there is a point in one cell within a distance $(2 + \Delta) r$ of some point in the other cell, where r is the transmission range of the nodes. By the definition of the Protocol Interference Model, if two cells are not interfering neighbors, transmissions in one cell do not interfere with transmissions in the other cell.

Lemma 1:   Each cell has no more than c interfering neighbors, where c is a constant that only depends on Δ.

Proof. We denote the length of each side of a cell (hexagon) as l and assume that $l = c_1 r$. Therefore, each cell is contained by a disk of radius $c_1 r$ and contains a disk of radius $\frac{\sqrt{3}}{2} c_1 r$.

By the definition of interfering neighbors, if a cell H' is an interfering neighbor of a cell H, one point in H' must be within a distance of $(2 + \Delta)\, r$ of a point in H. Figure 2 depicts the scenario where H' is the farthest interfering cell of H. For the Protocol Interference Model, all the interfering neighbors of H must be contained by a disk D of radius $3c_1 r + (2 + \Delta)r$. The number of interfering neighbors of a cell is the number of cells contained in disk D.

Since each cell contains a disk of radius $\frac{\sqrt{3}}{2}c_1 r$, the area of each cell is larger than the area of the contained disk, $\pi \left(\frac{\sqrt{3}}{2}c_1 r\right)^2$. The number of cells contained in disk D is upper bounded by

$$c = \pi \left((3c_1 + 2 + \Delta)\, r\right)^2 / \pi \left(\frac{\sqrt{3}}{2}c_1 r\right)^2 = \frac{4}{3} \left(\frac{3c_1 + 2 + \Delta}{c_1}\right)^2$$

Therefore, the number of interfering neighbors of a cell is upper bounded by c.

$\square$

Lemma 2: In the Protocol Model, there is a spatial scheduling policy that provides each cell one slot to transmit data in every $(c + 1)$ slots.

Proof. We construct the following graph. Each cell is represented by a vertex and edges are added between interfering neighbors. It is a well known fact in graph theory that a graph of degree no more than c can have its vertices colored by using no more than $(c + 1)$ colors, while no two adjacent vertices have the same color. Therefore, the cells in the network can be colored with no more than $(c + 1)$ colors, while no two interfering neighbors have the same color. We allow cells of the same color to transmit in the same time slot. These transmissions are from non-interfering cells, and thus, do not interfere with each other. Therefore, there exist scheduling schemes where each cell receives a slot for transmission every $(c + 1)$ slots. The degradation of network capacity due to the interference between the same type of traffic is bounded by a constant factor. $\square$

In the next two subsections, we derive the aggregate capacity of a wireless ad hoc network with regularly-placed base stations. The network capacity exhibits different capacities, depending on the asymptotic behavior of m as a function of n.

B. Capacity when $m = o(\sqrt{n})$

Theorem 2: The aggregate capacity of a wireless ad hoc network consisting of n nodes with an infrastructure network consisting of m regularly-placed base stations is

$$T(n, m) = \Theta \left(\sqrt{\frac{n}{\log \frac{n}{m^2}}} W_1\right) \tag{2}$$

under the deterministic routing strategy and channel allocation scheme when $m = o(\sqrt{n})$.

The proof is provided in Appendix II. From the derivation, we observe that the capacity contributed by ad hoc mode transmissions dominates as the number of wireless nodes, n, goes to infinity. The capacity contribution of infrastructure mode transmissions is negligible compared to that of ad hoc mode transmissions.

In this scenario, if the number of base stations is increased k times from m to km, the capacity of the network becomes

$$T(n, km) = \Theta\left(\sqrt{\frac{n}{\log \frac{n}{k^2 m^2}}} W\right) \simeq \Theta\left(\sqrt{\frac{n}{\log \frac{n}{m^2}}} W \left(1 + \frac{\log k}{\log \frac{n}{m^2}}\right)\right) = T(n, m)\left(1 + \frac{\log k}{\log \frac{n}{m^2}}\right)$$

In the above derivation, we used the fact that $\log \frac{n}{m^2} \gg 1$ since $n/m^2 \gg 1$. Therefore, a linear increase of the number of base stations only provides a logarithmic increase in the network capacity.

Now we investigate the benefit of infrastructure network to improve the capacity of a wireless ad hoc network. To quantify this benefit, we define the capacity gain factor as follows.

Definition 3: Capacity gain factor $g(n, m)$. The capacity gain factor of an infrastructure network consisting of m base stations is the ratio of the capacity of a wireless ad hoc network with infrastructure to the capacity of the wireless ad hoc network without infrastructure, i.e., $g(n, m) = T(n, m)/T_a(n)$, where $T_a(n)$ represents the aggregate capacity of a wireless ad hoc network of n nodes.

From [1], we know

$$T_a(n) = \Theta\left(\sqrt{\frac{n}{\log n}} W\right) \tag{3}$$

where W is the common transmission rate of the nodes over the wireless channel. The capacity gain factor is

$$g(n, m) = \Theta\left(\sqrt{\frac{\log n}{\log(n/m^2)}}\right) = \Theta\left(\frac{1}{\sqrt{1 - \frac{2\log m}{\log n}}}\right) \tag{4}$$

We now derive the capacity gain when the number of base stations scales as a polynomial of the number of nodes. Assuming $m = n^\alpha$ where $0 < \alpha < 1/2$, from Equation (4), we have

$$g(n, m) = \Theta\left(\frac{1}{\sqrt{1 - 2\alpha}}\right) \tag{5}$$

The gain factor is an increasing function of α, as shown in Fig. 3. If the number of base stations grows faster, the capacity gain is larger. Therefore, if m is bounded by two polynomials of n, n^{α_1} and n^{α_2}, where $\alpha_1 < \alpha_2$, the capacity gain is bounded by two constants:

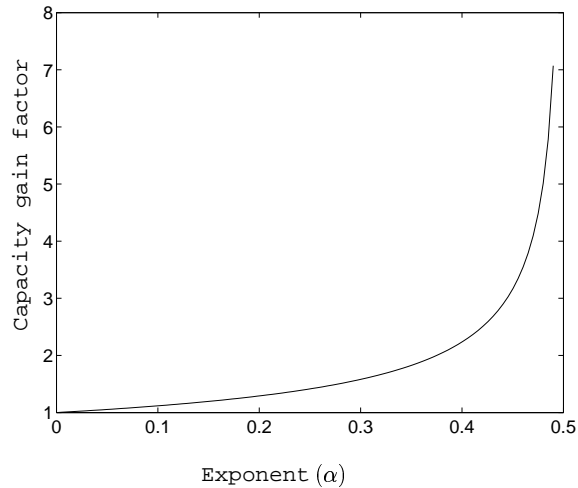$$\frac{c_2}{\sqrt{1 - 2\alpha_1}} \leq g(n, m) \leq \frac{c_2}{\sqrt{1 - 2\alpha_2}} \tag{6}$$

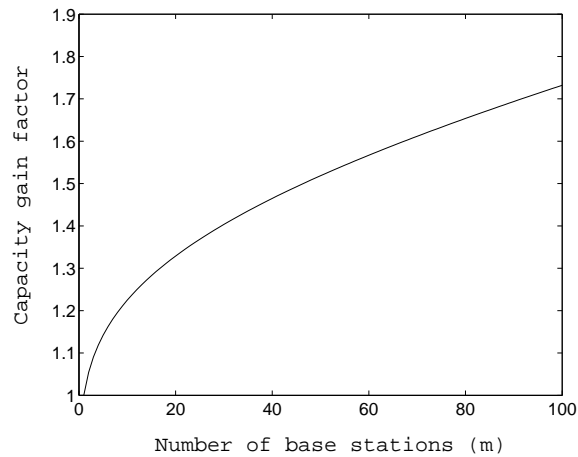Fig. 3.   Capacity gain factor as a function of α

Fig. 4.   Capacity gain factor as a function of number of base stations

where c_2 represents the constant factor in (4).

Figure 4 illustrates the benefit of placing more base stations to network capacity when the number of nodes in the network is fixed. We assume that there are 1,000,000 nodes and the number of base stations varies from 1 to 100. According to Equation (3), increasing the number of base stations from m to km provides less than a $\log k$-fold increase in the network capacity. The capacity of the ad hoc network does not change since n is fixed. As a result, the capacity gain factor only increases logarithmically with the number of base stations.

C. Capacity when $m = \Omega(\sqrt{n})$

In the previous case where $m = o(\sqrt{n})$, we have used the capacity result [1, Main Result 3] to obtain the per cell capacity contributed by ad hoc mode transmissions. However, this result is an asymptotic scaling law. It requires the number of nodes using ad hoc mode transmission to approach infinity. In this case, the number of nodes using ad hoc mode transmissions in a cell (n/m^2) is upper bounded by a constant. Hence, we cannot use [1, Main Result 3] to obtain the capacity contributed by ad hoc mode (T_a) as in the previous case. Instead, using Theorem 1, we can obtain the following theorem.

Theorem 3: The aggregate capacity of a wireless ad hoc network consisting of n nodes with an infrastructure network consisting of m regularly-placed base stations is

$$T(n, m) = \Theta(mW_2) \tag{7}$$

under the deterministic routing strategy and channel allocation scheme when $m = \Omega(\sqrt{n})$.

The proof of the theorem is presented in Appendix III. As the number of nodes, n, goes to infinity, the capacity contributed by infrastructure mode transmission dominates, while the contribution of ad hoc mode transmissions is negligible.

In this scenario, the capacity gain factor is

$$g(n, m) = \Theta\left(m\sqrt{\frac{\log n}{n}}\right) \tag{8}$$

Again, we are interested in the scaling behavior of capacity gain when m scales as a polynomial of n. Assume $m = n^\alpha$ where $1/2 \leq \alpha \leq 1$, a simple derivation yields

$$g(n, m) = \frac{1}{\sqrt{\alpha}} m^{1 - \frac{1}{2\alpha}} \sqrt{\log m} \tag{9}$$

If $\alpha = 1/2$, i.e., $m = \Theta(\sqrt{n})$, we have $g(n, m) = \Theta\left(\sqrt{\log m}\right)$. The capacity gain grows nearly logarithmically with the number of base stations. If $\alpha > 1/2$, $1 - \frac{1}{2\alpha} > 0$, the capacity gain grows polynomially with the number of base stations. In an extreme case where $\alpha = 1$, we have $m = \Theta(n)$, and $g = \Theta\left(\sqrt{m \log m}\right)$. The per node throughput capacity is

$$\lambda(n, m) = \Theta(1)$$

This suggests that if the number of base stations grows asymptotically at the same speed as the number of nodes, each node achieves constant throughput capacity. The reason is that in this case, each base station serves a constant number of nodes. Therefore, each node can connect to

the wired network using a constant share of the bandwidth, resulting in a $\Theta(1)$ per node capacity or a $\Theta(n)$ aggregate capacity.

Summarizing the above two subsections, the benefit of infrastructure network to improve network capacity depends on the relative number of base stations (m) to the number of wireless nodes (n). If m grows slower than $\sqrt{n}$, the capacity scales as $\Theta\left(\sqrt{\frac{n}{\log\frac{n}{m^2}}}\right)$. A linear increase of number of base stations only leads to a logarithmic increase of network capacity; the benefit of adding base stations is insignificant. However, if base stations can be added at a rate faster than $\sqrt{n}$, the capacity scales as $\Theta(m)$, which increases linearly with the number of base stations. In this case, the infrastructure network provides an effective improvement in the network capacity.

D. Effect of Random Deployment of Base Stations

All the above capacity results are obtained under the assumption that base stations are regularly deployed in a wireless ad hoc network. Now we study the capacity for the scenario where base stations are randomly and uniformly placed in a network, and compare the capacity to the case where base stations are regularly placed.

Theorem 4: For a wireless ad hoc network of n nodes with an infrastructure network of m randomly located base stations, if $m = o(\sqrt{n})$, the aggregate network capacity is $T(n, m) = \Theta\left(\sqrt{\frac{n}{\log\frac{n\log^2 m}{m^2}}}W_1\right)$; if $m = \Omega(\sqrt{n})$, the aggregate network capacity is $T(n, m) = \Theta\left(\frac{m}{\log m}W_2\right)$.

The proof of the theorem is provided in Appendix IV. As we can observe from the results for regularly-placed and randomly-placed base stations, when the deployment of base stations changes from regularly-placed to randomly-placed, the network capacity only suffers from a logrithmic degradation: when $m = o(\sqrt{n})$, the capacity decreases by a factor of $\log\log m$ from $\Theta\left(\sqrt{\frac{n}{\log\frac{n}{m^2}}}W_1\right)$ to $\Theta\left(\sqrt{\frac{n}{\log\frac{n\log^2 m}{m^2}}}W_1\right)$; when $m = \Omega(\sqrt{n})$, the network capacity decreases by a factor of $\log m$ from $\Theta(mW_2)$ to $\Theta\left(\frac{m}{\log m}W_2\right)$.

In this case, the logrithmic degradation is due to the irregularity of the base station locations. A cell now needs to be $\log m$ times larger than the regularly-placed case to ensure that there is a base station located inside.

In a recent work [15], the authors study the capacity of a wireless ad hoc network with randomly-deployed base stations, the same network model as being considered in this subsection. With the assumption that the number of base stations is on the same order as the number of wireless nodes $(m = \Theta(n))$, they show that the per-node capacity scales as $\Theta\left(\frac{1}{\log n}\right)$, a factor of

$\log n$ from the optimal value $\Theta(1)$. Note that the $m = \Theta(n)$ case considered in [15] is a special case of $m = \Omega(\sqrt{n})$, and the results they obtained are consistent with ours. If we let $m = \Theta(n)$, according to Theorem 4, the per-node capacity is $\Theta\left(\frac{1}{\log n}\right)$, the same as in [15]. Nevertheless, the authors of [15] prove the capacity is optimal among all possible routing strategies, while our results are for the deterministic routing strategy defined earlier.

VI. Conclusions

In this paper, we study the capacity of a wireless ad hoc network with infrastructure. Such a network architecture consists of a wireless ad hoc network and an infrastructure network of well-connected base stations. The goal of this work is to obtain the capacity of wireless ad hoc networks with infrastructure and investigate the benefit of infrastructure to improve network capacity.

We consider a wireless ad hoc network of n nodes with an infrastructure network of m base stations. For the deterministic routing strategy and regular placement of base stations, if m grows slower than $\sqrt{n}$, the capacity scales as $\Theta\left(\sqrt{\frac{n}{\log \frac{n}{m^2}}}\right)$. In this case, a linear increase of number of base stations only leads to a logarithmic increase of network capacity; the benefit of adding base stations is insignificant. However, if base stations can be added at a rate faster than $\sqrt{n}$, the capacity scales as $\Theta(m)$, which increases linearly with the number of base stations. The infrastructure network provides an effective improvement in the network capacity. Similar results are obtained for the case where base stations are randomly deployed.

Intuitively, the deterministic routing strategy improves network capacity in the following two ways. First, by using infrastructure network for sessions where destination nodes are not within the same cell as source nodes, the number of wireless transmissions is greatly reduced. The network can accommodate more data transmissions and hence the capacity is increases. Second, by using ad hoc mode for sessions where destinations are within the same cell as sources, the available spatial concurrency is utilized, which further increases the network capacity. Although we have shown that the deterministic routing strategy leads to an improved network capacity, it is still not clear if it is optimal among all possible routing strategies from the network capacity perspective.

In addition to the deterministic routing strategy, we also considered a probabilistic routing strategy in [18], where a node probabilistically chooses whether or not to use infrastructure to send data. Similar results are obtained for network capacity under the probabilistic routing strategy. We show that if m grows slower than $\sqrt{\frac{n}{\log n}}$, the capacity has the same asymptotic behavior as a pure ad hoc network. In this case, the probabilistic routing strategy yields a smaller capacity

than the deterministic routing strategy. There is no benefit in using the infrastructure. However, if m grows at a rate faster than $\sqrt{\frac{n}{\log n}}$, the capacity scales as $\Theta(m)$, which is the same as under the deterministic routing strategy. In this case, the infrastructure network provides an effective improvement in capacity over a pure ad hoc network.

Therefore, for both routing strategies, in order to achieve a non-negligible capacity gain, the investment in the wired base station infrastructure should be high enough: base stations should be added at a rate faster than $\sqrt{n}$ for the deterministic routing strategy and $\sqrt{\frac{n}{\log n}}$ for the probabilistic routing strategy.

APPENDIX I

PROOF OF THEOREM 1

Without loss of generality, we assume $s \geq d$, i.e., $min(s, d) = d$. The case where $s < d$ can be established in a similar way. Also, in the following proof, we assume the source node set and destination node set are two disjoint sets. In the general case where there is overlap between the source set and destination set, the capacity is decreased at most by a factor of two, which can be achieved as follows. For each node that is both a source and destination, we let it alternate between the source and destination roles, each lasting one time slot.

We divide the possible values of d into two different ranges and prove the theorem for each of the two scenarios: (i) $d = O\left(\sqrt{\frac{k}{\log k}}\right)$, and (ii) $d = \omega\left(\sqrt{\frac{k}{\log k}}\right)$.

Scenario (i): $d = O\left(\sqrt{\frac{k}{\log k}}\right)$.

In this scenario, we need to prove $T_a(k, s, d) = \Theta(dW)$. Since the maximum transmission rate of a node over the common wireless channel is W bits/sec and a node can only receive data from one sender at the same time, the aggregate capacity of the network is upper bounded by dW.

Now we present a constructive lower bound on the aggregate capacity. We first derive the capacity of the case where only d out of s source nodes are active sending data, $T_a(k, d, d)$. We then prove $T_a(k, s, d) \geq T_a(k, d, d)$ for $d \leq s$. Thus the lower bound for $T_a(k, d, d)$ is necessarily a lower bound for $T_a(k, s, d)$. The derivation of $T_a(k, d, d)$ is very similar to [1, Section 4]. For completeness, we present the key steps of the constructive proof and highlight the differences in our case. As in [1], the capacity result is first derived for a sphere S^2 case, which can then be proved to hold for a planar disk. Steps 1 through 5 are essentially the same as in [1].

Step 1: Sphere S^2 contains k randomly located nodes. We partition the sphere into a Voronoi tessellation (V_k) such that each Voronoi cell contains a disk of radius $\rho(k)$, and each cell is

contained by a disk of radius $2\rho(k)$, where

$$\rho(k) := \text{radius of a disk of area } \frac{100 \log k}{k} \text{ on } S^2$$

Step 2: Choose a transmission range $r(k) = 8\rho(k)$, so every node in a cell is within a distance $r(k)$ from every node in its own or adjacent cell.

Step 3: Every cell in V_k has no more than a constant number (c_3) of interfering neighbors. Therefore, there is a spatial scheduling policy such that each cell gets one slot to transmit data in every constant number $(1 + c_3)$ of time slots.

Step 4: Denote L_i as the straight line segment connecting S_i and D_i, where S_i is the location of source i and D_i is the location of the corresponding destination. The random sequence of straight line segments $\{L_i\}_{i=1}^d$ form an i.i.d. sequence. Thus allows us to apply the law of large numbers later.

The routing strategy is the same as in [1]. The route of a packet approximates the straight line segment between the source and destination pair. More specifically, a packet is forwarded in a multi-hop fashion, and in each hop, a packet is forwarded from one cell to another in the order in which the cells intersect the line from source to destination.

Step 5: It can be proved that, with probability 1, each cell in V_k contains at least one node as $k \to \infty$. Therefore, the above routing strategy is feasible, since each cell contains at least a node to relay traffic with a transmission range enough to reach all nodes in any adjacent cells.

Step 6: This step differs from [1, Section 4.8, 4.9].

Lemma 4.9 in [1] hold without modification. For every line L_i and cell $V \in V_k$,

$$Prob(\text{Line } L_i \text{ intersects } V) \leq c_4 \sqrt{\frac{\log k}{k}}$$

Unlike in [1], there are only d lines in $\{L_i\}_{i=1}^d$ in this case, where $d = O\left(\sqrt{\frac{k}{\log k}}\right)$. Thus, the mean number of lines passing through a cell in bounded by a constant.

$$E[\text{Number of lines in } \{L_i\}_{i=1}^d \text{ intersecting a cell } V] \leq c_5$$

Therefore, the actual traffic served by each cell is bounded by a constant. There is a $\delta(k) \to 0$ such that

$$Prob\left(\sup_{V \in V_k} (\text{traffic needing to be carried by cell } V) \leq c_5 \lambda(k, d, d)\right) \geq 1 - \delta(k)$$

where $\lambda(k, d, d)$ is the per-node capacity.

Step 7: From step 3, the rate at which each cell can transmit is $\frac{W}{1+c_3}$. From step 6, the rate at which each cell needs to transmit is less than $c_5\lambda(k, d, d)$. This rate is achievable if it is less than the available rate.

$$c_5\lambda(k, d, d) \leq \frac{W}{1 + c_3}$$

Therefore, the aggregate capacity is lower bounded by

$$T(k, d, d) = d\lambda(k, d, d) = \Theta(dW)$$

Step 8: We now need to prove $T_a(k, s, d) \geq T_a(k, d, d)$ for $s \geq d$. For an ad hoc network of k nodes, if a capacity of $T_a(k, d, d)$ can be achieved when there are d source nodes and d destination nodes, at least the same capacity can be achieved if there are $s > d$ source nodes and d destination nodes. The reason is as follows. At each time slot, we randomly choose d sources from the source set (containing s source nodes) for transmission. The other sources are refrained from sending data. This policy necessarily achieves capacity $T_a(k, d, d)$. Therefore, $T_a(k, s, d)$ is greater than or equal to $T_a(k, d, d)$, and hence, lower bounded by $\Theta(dW)$.

We have proved $T_a(k, s, d)$ is upper and lower bounded by dW. Therefore,

$$T_a(k, s, d) = \Theta(dW) = \Theta(min(s, d)W)$$

Scenario (ii): $d = \omega\left(\sqrt{\frac{k}{\log k}}\right)$.

In this scenario, we need to prove $T_a(k, s, d) = \Theta\left(\sqrt{\frac{k}{\log k}}W\right)$.

For the upper bound, using similar arguments as in step 8 above, we have $T_a(k, k, k) \geq T_a(k, s, k) \geq T_a(k, s, d)$. If a network can support an aggregate capacity of $T_a(k, s, d)$ when there are s source nodes and d destination nodes, at least the same capacity can be supported by the network when there are k source nodes and k destination nodes. This is true if the additional $(k - s)$ source nodes and $(k - d)$ destination nodes do not send and receive data traffic. From [1], $T_a(k, k, k)$ is upper bounded by $O\left(\sqrt{\frac{k}{\log k}}\right)$, therefore, so is $T_a(k, s, d)$.

To obtain the lower bound for $T_a(k, s, d)$, we first derive a lower bound for $T_a(k, d, d)$, which provides a lower bound for $T_a(k, s, d)$. The steps are similar to those in case (i). Here we highlight the differences.

Steps 1 through 5 are the same as above.

Step 6: Since there is a total number of d lines in $\{L_i\}_{i=1}^d$, where $d = \omega\left(\sqrt{\frac{k}{\log k}}\right)$, the mean number of lines passing through a cell in bounded by

$$E[\text{Number of lines in } \{L_i\}_{i=1}^d \text{ intersecting a cell } V] \leq c_6 d\sqrt{\frac{\log k}{k}}$$

Therefore, there is a $\delta(k) \to 0$ such that

$$Prob\left(\sup_{V \in V_k}(\text{traffic needing to be carried by cell } V) \leq c_6\lambda(k, d, d)d\sqrt{\frac{\log k}{k}}\right) \geq 1 - \delta(k)$$

where $\lambda(k, d, d)$ is the per-node capacity in this case.

Step 7: The rate at which each cell can transmit is $\frac{W}{1+c_3}$. From step 6, the rate at which each cell needs to transmit is less than $c_6\lambda(k, d, d)d\sqrt{\frac{\log k}{k}}$. This rate is achievable if it is less than the available rate.

$$c_6\lambda(k, d, d)d\sqrt{\frac{\log k}{k}} \leq \frac{W}{1 + c_3}$$

Therefore, the aggregate capacity is

$$T(k, d, d) = d\lambda(k, d, d) \leq \sqrt{\frac{k}{\log k}}\frac{W}{c_6(1 + c_3)}$$

Step 8: This is the same as in case (i).

Therefore, we have proved that $T_a(k, s, d)$ is upper and lower bounded by $\Theta\left(\sqrt{\frac{k}{\log k}}W\right)$.

Summarizing (i) and (ii), the capacity is given by

$$T(k, s, d) = \Theta\left[\min\left(\min(s, d)W, \sqrt{\frac{k}{\log k}}W\right)\right]$$

$\square$

APPENDIX II

PROOF OF THEOREM 2

To derive the aggregate network capacity, we first compute the per-cell capacity contributed by ad hoc mode transmissions (T_a) and by infrastructure mode transmissions (T_i), respectively. The aggregate capacity of the network can be computed based on the per-cell capacity and Lemma 2.

Consider an arbitrary cell h, let Y_i be a random variable that represents whether node i ($1 \leq i \leq n$) and its destination are both located in cell h. The random variables are defined as follows:

$$Y_i = \begin{cases} 1 & \text{both node } i \text{ and its destination are in cell } h \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

In a wireless ad hoc network with an infrastructure network of m base stations, there are m cells. Nodes and the corresponding destinations are randomly and independently and placed in the unit area disk. The probability that a node i is located in cell h is $1/m$; the probability that the destination of node i is located in cell h is also $1/m$. Therefore, $E[Y_i] = 1/m^2$.

We then introduce a random variable $N_h = \sum_{i=1}^n Y_i$, representing the number of source and destination pairs communicating using ad hoc mode within cell h. Since $\{Y_i\}_{i=1}^n$ is an i.i.d. sequence of random variables with $E[Y_i] = 1/m^2$. By the Law of Large Numbers, with probability 1,

$$\frac{N_h}{n} = \frac{1}{n} \sum_{i=1}^n Y_i \to \frac{1}{m^2} \text{ as } n \to \infty \tag{11}$$

Given $m = o(\sqrt{n})$, we have $\lim_{n\to\infty} n/m^2 \to \infty$, and thus $\lim_{n\to\infty} N_h \to \infty$. According to [1], for a random ad hoc network of N_h nodes and a common transmission rate of W_1, the per node capacity is $\Theta\left(\frac{W_1}{\sqrt{N_h \log N_h}}\right)$, as N_h goes to infinity. Therefore, the capacity of cell h contributed by ad hoc transmissions is $T_a(N_h) = \Theta\left(\sqrt{\frac{N_h}{\log N_h}} W_1\right)$. Denote

$$c_7 = \lim_{n\to\infty} \inf \frac{T_a(N_h)}{\sqrt{\frac{N_h}{\log N_h}} W_1}, \text{ and } c_8 = \lim_{n\to\infty} \sup \frac{T_a(N_h)}{\sqrt{\frac{N_h}{\log N_h}} W_1}$$

By (11), we have $\lim_{n\to\infty} \frac{\sqrt{\frac{N_h}{\log N_h}}}{\sqrt{\frac{n/m^2}{\log(n/m^2)}}} = 1$. Therefore,

$$\lim_{n\to\infty} \inf \frac{T_a(N_h)}{\sqrt{\frac{n/m^2}{\log(n/m^2)}} W_1} = \lim_{n\to\infty} \inf \frac{T_a(N_h)}{\sqrt{\frac{N_h}{\log N_h}} W_1} \frac{\sqrt{\frac{N_h}{\log N_h}}}{\sqrt{\frac{n/m^2}{\log(n/m^2)}}} = c_7$$

$$\lim_{n\to\infty} \sup \frac{T_a(N_h)}{\sqrt{\frac{n/m^2}{\log(n/m^2)}} W_1} = \lim_{n\to\infty} \sup \frac{T_a(N_h)}{\sqrt{\frac{N_h}{\log N_h}} W_1} \frac{\sqrt{\frac{N_h}{\log N_h}}}{\sqrt{\frac{n/m^2}{\log(n/m^2)}}} = c_8$$

The term $\lim_{n\to\infty} T_a(N_h)/\sqrt{\frac{n/m^2}{\log(n/m^2)}} W_1$ is upper and lower bounded by two constants. Therefore, the per-cell capacity contributed by the ad hoc mode communications is

$$T_a = \Theta\left(\sqrt{\frac{\frac{n}{m^2}}{\log \frac{n}{m^2}}} W_1\right) \tag{12}$$

Now we calculate the capacity contributed by infrastructure mode communication. Since the same packet has to go through an uplink and a downlink when transmitted in infrastructure mode, it should be only counted once in the capacity. Recall that the bandwidths assigned to carry uplink and downlink traffic are W_2 and W_3, and we have $W_2 = W_3$.

Since all the infrastructure mode traffic has to go through the base station and the base station can only receive data at the rate of W_2 bits/sec at any time. T_i is upper bounded by W_2. For the lower bound, if each node in the infrastructure mode employs a transmission range of l (the side

length of each cell), there is a schedule for each node to transmit to the base station in a round robin fashion, yielding a capacity of W_2. Therefore,

$$T_i = \Theta(W_2). \tag{13}$$

Since there is no interference between the ad hoc mode and the infrastructure mode, the capacity of a cell is just $T_a + T_i$.

Moreover, in Lemma 2, we proved that there is a scheduling policy such that each cell gets a slot to transmit in every constant number $(c+1)$ of time slots. Therefore, the aggregate capacity of the network is

$$T(n,m) = \Theta\left(mT_a + mT_i\right) = \Theta\left(\sqrt{\frac{n}{\log\frac{n}{m^2}}}W_1 + mW_2\right) = \Theta\left(\sqrt{\frac{n}{\log\frac{n}{m^2}}}W_1\right)$$

In the last step, we used the fact that $\lim_{n\to\infty} m/\sqrt{\frac{n}{\log\frac{n}{m^2}}} = 0$, which indicates that the capacity contributed by ad hoc mode transmissions dominates as n goes to infinity. The capacity contribution of infrastructure mode transmissions is negligible compared to that of ad hoc mode transmissions.

$\square$

APPENDIX III

PROOF OF THEOREM 3

The value of n/m^2 does not change the capacity contributed by the infrastructure mode. We have $T_i = \Theta\left(W_2\right)$.

When $m = \Omega(\sqrt{n})$, the number of nodes communicating in ad hoc mode, n/m^2, is upper bounded by a constant, we cannot use [1, Main Result 3] to derive T_a, since it requires that the number of nodes using ad hoc mode transmissions approaches infinity. Instead, we can obtain T_a using Theorem 1.

Consider a cell h in a wireless ad hoc network of n nodes with m base stations, by the Law of Large Numbers, with probability 1, the number of nodes in cell h approaches n/m as $n \to \infty$. From Equation (11), we know that with probability 1, the number of source and destination nodes communicating using the ad hoc mode within cell h approaches n/m^2 as $n \to \infty$. Using the notations in Theorem 1, we have $k = n/m$, and $s = d = n/m^2$. Since $m = \Omega(\sqrt{n})$, we have $s = d = O\left(\sqrt{\frac{k}{\log k}}\right)$. Therefore, using Theorem 1, we obtain

$$T_a = \Theta\left(\frac{n}{m^2}W_1\right)$$

Hence, the aggregate capacity of the network is

$$T(n, m) = \Theta\left(mT_a + mT_i\right) = \Theta(\frac{n}{m}W_1 + mW_2) = \Theta(mW_2)$$

In the derivation, we used the fact $m = \Omega(n/m)$, which indicates that as n goes to infinity, the capacity contributed by infrastructure mode transmission dominates, while the contribution of ad hoc mode transmissions is negligible.

$\square$

APPENDIX IV

PROOF OF THEOREM 4

We divide the space into regular "virtual" cells of area $16 \log m/m$. There are a total number of $m/16 \log m$ such cells. We first prove that with high probability, each virtual cell contains at least one base station; and the number of empty cells, i.e., cells that do not have a base station, is zero.

Consider an arbitrary virtual cell j, let $X_i (1 \leq i \leq m)$ be a random variable indicating whether base station i is located inside cell j.

$$X_i^j = \begin{cases} 1 & \text{base station } i \text{ is located inside cell } j \\ 0 & \text{otherwise} \end{cases}$$

We define random variable $Y_j = \sum_{i=1}^{m} X_i^j$ to represent the total number of base stations that are located in cell j. Since the base stations are randomly and uniformly distributed in the network, the probability that any given base station is located inside cell j is $16 \log m/m$, and the expected number of base stations located in the cell is

$$\mu_j = E[Y_j] = 16 \log m$$

Using a standard Chernoff bound argument, $\forall \delta \in (0, 1]$,

$$Prob(Y_j > (1 - \delta)\mu_j) > 1 - exp(-\mu_j \delta^2/2)$$

Let $\delta_j = 1/2$, we have

$$Prob(Y_j > \frac{1}{2} \log m) > 1 - 1/m^2 \tag{14}$$

Therefore, each virtual cell contains at least one base station with high probabilty. Now let N denote the number of empty cells, $N = \sum_j \mathbf{1}\{Y_j = 0\}$, where $\mathbf{1}\{Y_j = 0\}$ is the indicator function.

Based on (14), we have

$$Prob(Y_j < 1) < Prob(Y_j < \frac{1}{2}\log m) < 1/m^2$$

Hence,

$$Prob(N > 0) < \sum_j Prob(Y_j < 1) < m/(16 \log m) Prob(Y_j < 1) < 1/(16m \log m)$$

or equivalently,

$$Prob(N = 0) = 1 - Prob(N > 0) > 1 - 1/(16m \log m)$$

Therefore, the number of empty cells equals to zero with high probability.

We adopt the following routing strategy similar to what we have used for the regular network scenario: if the destination node is located in the same virtual cell as the source node, data are forwarded in ad hoc mode. Otherwise, if the destination node is not located in the same cell as the source node, data are forwarded in infrastructure mode.

Now we have a virtual regular network of n nodes and $m/\log m$ cells. Using the same technique as in the proofs of Theorem 2 and Theorem 3, we can obtain the results in Theorem 4.

□

REFERENCES

[1] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, Mar 2000.

[2] Y.-D. Lin and Y.-C. Hsu, "Multihop cellular: A new architecture for wireless communications," in *Proc. IEEE Infocom*, 2000.

[3] O. Dousse, P. Thiran, and M. Hasler, "Connectivity in ad-hoc and hybrid networks," in *Proc. IEEE Infocom*, 2002.

[4] H. Luo, R. Ramjee, P. Sinha, L. Li, and S. Lu, "Ucan: A unified cellular and ad-hoc network architecture," in *Proc. of MobiCom*, 2003.

[5] S. De, O. Tonguz, H. Wu, and C. Qiao, "Integrated cellular and ad hoc relay (icar) systems: Pushing the performance limits of conventional wireless networks," in *Proc. Hawaii International Conference on System Sciences*, 2002.

[6] H.-Y. Hsieh and R. Sivakumar, "On using the ad-hoc network model in cellular packet data networks," in *Proc. of MobiHoc*, 2002.

[7] P. Gupta and P. Kumar, "Internets in the sky: The capacity of three dimensional wireless networks," *Communications in Information and Systems*, vol. 1, no. 1, pp. 39–49, 2001.

[8] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad-hoc wireless networks," in *Proc. IEEE Infocom*, 2001.

[9] N. Bansal and Z. Liu, "Capacity, mobility and delay in wireless ad hoc networks," in *Proc. IEEE Infocom*, 2003.

[10] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: the relay case," in *Proc. IEEE Infocom*, 2002.

[11] D. Marco, E. Duarte-Melo, M. Liu, and D. L. Neuhoff, "On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data," in *Proc. International Workshop on Information Processing in Sensor Networks (IPSN)*, 2003.

[12] J. Li, C. Blake, D. S. J. D. Couto, H. I. Lee, and R. Morris, "Capacity of ad hoc wireless networks," in *Mobile Computing and Networking*, 2001, pp. 61–69.

[13] C. Peraki and S. Servetto, "On the maximum stable throughput problem in random networks with directional antennas," in *Proc. of The Fourth ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2003.

[14] S. Yi, Y. Pei, and S. Kalyanaraman, "On the capacity improvement of ad hoc wireless networks using directional antennas," in *Proc. of The Fourth ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2003.

[15] U. Kozat and L. Tassiulas, "Throughput capacity of random ad hoc networks with infrastructure support," in *Proc. of MobiCom*, 2003.

[16] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*. MIT Press, 1990.

[17] B. Liu, Z. Liu, and D. Towsley, "Capacity of a wireless ad hoc network with infrastructure," Computer Science Dept - University of Massachusetts Amherst, Tech. Rep., 2004.

[18] ——, "On the capacity of hybrid wireless networks," in *Proc. IEEE Infocom*, 2003.