WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2004; **4**:263–279 (DOI: 10.1002/wcm.216)

On the scalability and capacity of planar wireless networks with omnidirectional antennas

Onur Arpacioglu*,† and Zygmunt J. Haas

School of Electrical and Computer Engineering, Wireless Networks Laboratory, Cornell University, Ithaca, NY 14853, U.S.A.

Summary

We extend the previously well-known results on the capacity of wireless networks and present the implications of our results on network scalability. In particular, we find bounds on the maximum achievable per-node end-to-end throughput, λ_e, and the maximum number of simultaneously successful wireless transmissions, N_t^{max}, under a more general network scenario than previously considered. Furthermore, in the derivation of our results, we make no restrictions on the mobility pattern of the nodes or on the number simultaneous transmissions and/or receptions that nodes are capable of maintaining. In our derivation, we analyze the effect of parameters such as the area of the network domain, A, the path loss exponent, γ, the processing gain, G, and the *SINR* threshold, β. Specifically, we prove the following results for a wireless network of N nodes that are equipped with omnidirectional antennas:

(1) λ_e is $\Theta(1/N)$ under very general conditions. This result continues to hold even when the communication bandwidth is divided into sub-channels of smaller bandwidth.
(2) N_t^{max} has an upper bound that does not depend on N, which is the simultaneous transmission capacity of the network domain, $N_t^{\mathbf{Q}}$. For a circular network domain, $N_t^{\mathbf{Q}}$ is $O(A^{\min\{\gamma/2,1\}})$ if $\gamma \neq 2$ and $O(A/\log(A))$ if $\gamma = 2$. In addition, $N_t^{\mathbf{Q}}$ is $O(\gamma^2)$ and $O(G/\beta)$. Moreover, lack of attenuation and lack of space are equivalent, where $N_t^{\mathbf{Q}}$ cannot exceed $1 + G/\beta$.
(3) As $N \to \infty$ a desired per-node end-to-end throughput is not achievable, unless the average number of hops between a source and a destination does not grow indefinitely with N, A grows with N and N is $O(A^{\min\{\gamma/2,1\}})$ if $\gamma \neq 2$ and $O(A/\log(A))$ if $\gamma = 2$. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: network scalability; network capacity; network throughput; omnidirectional transmission

1. Introduction

Scalability of wireless networks has been an important research topic in the recent years, because of the growing demand to support a large number of nodes in certain types of wireless networks such as sensor networks, which can potentially consist of millions of nodes. Two important questions in this context are:

*Correspondence to: Onur Arpacioglu, School of Electrical and Computer Engineering, Wireless Networks Laboratory, Cornell University, Ithaca, NY 14853, U.S.A.
†E-mail: o.arpacioglu@cornell.edu

Contract/grant sponsor: Office of Naval Research (DoD Multidisciplinary University Research Initiative (MURI) program); contract/grant number: N00014-00-1-0564.
Contract/grant sponsor: Air Force Office of Scientific Research (DoD MURI program); contract/grant number: F49620-02-1-0233.
Contract/grant sponsor: National Science Foundation; contract/grant number: ANI-0081357.

Copyright © 2004 John Wiley & Sons, Ltd.

(1) are wireless networks scalable? In other words, is it possible to support a large number of users in a wireless network? (2) If there are scalable patterns of wireless networks, what are the conditions that govern their scalability?

To answer these questions, one must first define what scalability is. In this work, we interpret scalability as: non-vanishing per-node end-to-end throughput, bounded end-to-end delay, bounded power consumption, bounded processing power and bounded memory requirement at each node, as the number of nodes grows large.

In this paper, we focus on the throughput aspect of scalability. Our objectives are: (1) to obtain theoretical results that show the dependencies among the per-node end-to-end throughput capacity λ_e, N and other parameters of a wireless network and (2) to determine the implications of these results on the scalability of the wireless network.

One of the most well-known works on capacity of wireless networks was published by Gupta and Kumar [1]. Their work stimulated the scientific community to search for a better understanding of what are the scalability bounds of such networks. In that paper, a theoretical framework to analyze the capacity of peer-to-peer wireless networks was formalized through two network models. The first network model, the *arbitrary network* model, assumes that all N nodes in the network are static, there are no restrictions on nodes' locations and the network domain (i.e. the region within which the nodes are located) is a circular disk of area 1 m^2. Each node is capable of maintaining at most one transmission or one reception at any given time. There are no restrictions on the choice of transmission powers, traffic pattern, routing protocol and spatial-temporal transmission scheduling policy. The second model is the *random network* model, which assumes a uniform distribution of node locations, a random traffic pattern and a fixed transmission power that is selected to ensure the connectivity of the network as N tends to infinity.

Additionally, in Reference [1], two models for successful reception are proposed. The first reception model is the protocol model, which considers a transmission as unsuccessful if the receiver is within the interfering range of an unintended transmitter. The second model is the physical model, which better represents realistic reception in practical wireless networks. In the physical model, for a transmission to be successful, the signal-to-interference-and-noise ratio, SINR, at a receiver has to be above some threshold value. It is assumed that the antennas are omnidirectional and that P/x^γ is the power received at a distance x from a transmitter, where P is the transmitted power and the path loss exponent γ is assumed to be larger than 2.

Reference [1] concluded that, with the protocol model, λ_e is $\Theta(1/\sqrt{N})$ for arbitrary networks, whereas λ_e is $\Theta\left(1/\sqrt{N\log(N)}\right)$ for random networks. With the physical model, they concluded that λ_e is $O(1/N^{1/\gamma})$ and $\Omega(1/\sqrt{N})$ for arbitrary networks, whereas, λ_e is $O(1/\sqrt{N})$ and $\Omega\left(1/\sqrt{N\log(N)}\right)$ for random networks.

One of the assumptions of Reference [1] is that all nodes are immobile. Grossglauser and Tse [2], explored whether or not introducing mobility can increase λ_e. Their network model introduced some additional restrictions on the random network model of Reference [1]. Firstly, they used the physical model, but allowed wideband communication by incorporating the processing gain, as to reduce interference. Secondly, the locations of the mobile nodes form a stationary ergodic process with a uniform stationary distribution in the network domain. Thirdly, source–destination pairs do not change. Finally, they assumed that very long end-to-end packet delays are tolerable. Grossglauser and Tse [2], concluded that there exists a routing and scheduling policy that delivers a packet to its destination with not more than two hops and allows λ_e to be $\Theta(1)$ as N becomes large. Moreover, both References [1] and [2] concluded that it is possible to schedule $\Theta(N)$ many simultaneously successful transmissions in a wireless network.

Toumpis and Goldsmith [3], used an SINR dependent rate model. Using numerical methods, they evaluated the effect of spatial reuse, multi-hop routing, power control and successive interference cancellation for a particular placement of nodes. They concluded that each of these schemes provides expansions of the capacity region that is defined by the set of achievable rates between the nodes.

Li *et al.* [4], pointed out the effect of traffic pattern on λ_e and concluded that only wireless networks with local traffic patterns can be scalable.

Usage of directional antennas at the transmitters or the receivers can provide significant increases in λ_e, depending on how narrow the width of the antenna's main lobe and how small the side lobes of the antenna radiation pattern can be made. For example, using a sender-based interference model, Yi, Pei and Kalyanaraman [5], investigated the improvement in

Copyright © 2004 John Wiley & Sons, Ltd.

λ_e provided by the usage of directional antennas for arbitrary and random wireless networks.

The ability of a node to maintain multiple simultaneous transmissions and/or receptions can also provide an increase in λ_e. For example using directional antennas, Peraki and Servetto [6], studied random networks with and without multiple simultaneous transmission or reception capability and concluded that an improvement of at most $\Theta(\log^2(N))$ can be achieved over the $\Theta\left(1/\sqrt{N\log(N)}\right)$ result of Reference [1].

Deployment of a wired backbone can also provide an increase in λ_e, since it allows reducing average number of wireless transmissions per packet. For example, Liu, Liu and Towsley [7], considered the benefit of deploying base stations connected to a wired backbone in the random network of Reference [1]. They concluded that if the number of base stations grows asymptotically faster than $\sqrt{N}$, then aggregate throughput capacity increases linearly with the number of base stations.

There have also been information theoretical approaches to analyze the throughput capacity problem, such as References [8] and [9]. These approaches concluded that sizable gains in network throughput can be provided by not treating interference simply as noise but rather employing more sophisticated receivers and implementing certain cooperation strategies among the nodes during transmissions.

Our work has been motivated by the desire to relax some of the limitations of References [1] and [2], and to improve on their models. In particular, the radio propagation model (P/x^{γ}), which has been widely used in such studies, becomes inappropriate as the transmitter–receiver separation becomes small. Since the network domain in these studies is a disk of area $1\,\mathrm{m}^2$, the model gives unrealistic results when the nodal density increases beyond some limit. In terms of mobility, in Reference [1] nodes are immobile, while in Reference [2] the mobility pattern is a very special one. In our work, we allow for a general mobility pattern of the nodes. Furthermore, we relax the assumptions in Reference [2] that the source–destination pairs never change and that the end-to-end packet delays can be unbounded. Moreover, in References [1] and [2] each node can maintain either a single transmission or reception at a given time, whereas our work also considers the situation when the nodes can maintain multiple transmissions and/or receptions at the same time. In addition, we also analyze the effect of parameters such as A, γ, G and β on λ_e, none of which has been addressed in previous works.

The main contribution of this study is in the derivation of a new approach to analyze the scalability patterns of wireless networks through the use of a more general network model as compared with the network models of previous studies.

The rest of the paper is organized as follows: Section 2 presents our network model and Section 3 provides related definitions. In Section 4, we derive the upper bounds on the simultaneous transmission capacity and the per-node end-to-end throughput capacity. In Section 5, we analyze the derived upper bounds, together with illustrative figures. Section 6 completes the proof of 'λ_e is $\Theta(1/N)$'. Section 7 discusses the implications of our results on scalability. Finally, we conclude in Section 8.

2. Network Model

In this section, we explain the assumptions underlying our results. Some of the assumptions are quite general, allowing our upper bound results to hold even in networks that are configured with optimal settings of the parameters. We will elaborate on this further, after presenting our results.

2.1. Network Domain and Nodes

Network domain is defined to be the space within which each node is constrained to reside. We denote this space by $\boldsymbol{Q}$. We will assume that $\boldsymbol{Q}$ is a closed disk with a diameter D and an area A. There are N nodes in the network domain and no restrictions on the mobility pattern of the nodes within $\boldsymbol{Q}$.

2.2. Transmitter and Receiver Model

Each of the nodes is capable of being a transmitter and/or a receiver at any given time. All transmitters and receivers are equipped with omnidirectional antennas. There are no restrictions on the variation of transmission power during a transmission or on the number of simultaneous transmissions and/or receptions that a node is capable of maintaining. Hence, the usual assumption that a node is capable of maintaining either one transmission or one reception at any given time is one of the many cases covered by our model. For the time being, we assume that all transmissions take place within the

Copyright © 2004 John Wiley & Sons, Ltd.

same communication bandwidth but we will relax this assumption later. At the intended receiver of a transmitted signal, all of the remaining received signals are considered as interference. $\zeta_i(t)$ is the power of the thermal noise present in the communication bandwidth at receiver i at time t. We assume that each of the receivers can receive information intended for itself reliably at a rate not larger than $W_{\max}$ bits/s and only when the SINR at the receiver is greater than or equal to the SINR threshold, β. Information received when the above condition does not hold is considered unreliable and, thus discarded. In this paper, we let β be any positive real number. In general, β is dependent on the modulation scheme, the required bit error rate of the received information, the required transmission rate and the type of the error control code. The processing gain, G, is the factor by which the total received interference power is reduced at each of the receivers. In this work, we let G be any positive real number. Typically $G > 1$ for wideband communication systems, such as spread spectrum CDMA, and is taken to be 1 for narrowband communications.

2.3. Propagation Model

Let $P_{\rm t}^j(t)$ be the power transmitted by the transmitter j at time t. Let $d_{ji}(t)$ be the distance[‡] between the transmitter j and the receiver i at time t. Let $P_r^{ji}(t)$ be the power received by the receiver i from the transmitter j at time t. We will assume that

$$P_r^{ji}(t) = P_{\rm t}^j(t)a(d_{ji}(t)), \tag{a.1}$$

where $a(x)$ is the *attenuation function*. One of the most commonly used expressions for $a(x)$ is $1/x^\gamma$, where $\gamma \geq 0$ is the path loss exponent.[§] However, this expression becomes inappropriate as x becomes smaller than 1, as it results in receiving power that is larger than the transmitted power. In fact, the received power in the formula approaches infinity as x tends to zero. Of course, this is unrealistic and is an artifact of the inappropriateness of the expression for small distances.[¶] The inappropriateness of the $1/x^\gamma$ formula was also noticed in some previous works and to obtain more meaningful results at small distances, while approximating the conventional model at large distances, the following alternative propagation model was proposed in those studies (see, e.g. [12] and [13]):

$$a(x) = \frac{1}{(1+x)^\gamma}, \quad x \geq 0. \tag{a.2}$$

We call this propagation model, the *power law decaying propagation model*[||] and we will use this model in our calculations.

2.4. Traffic Pattern

In this work, there are no restrictions on the temporal variation of destination of each of the information sources, the route of the information through the network and the segmentation of information, so that different segments can possibly be transmitted over different paths and at different times. As in References [1] and [2], we assume that intermediate nodes do not jointly encode and transmit information

[‡]In this paper, all distance measures and area measures are in units of 'm' and 'm^2' respectively.

[§]In free space $\gamma = 2$, but in realistic mobile radio channels, γ can take values between 1.6 and 6 (see [10] and [11]).

[¶]The precise reason for this problem can be explained as follows: consider the free space case. In the derivation of the received power expression, a unity gain point source in free space is assumed and the flux of the transmitted power $P_{\rm t}$, per unit surface area of the sphere with radius x around the source is calculated. The resulting power flux density expression, $P_{\rm t}/(4\pi x^2)$, has the unit watts/meter^2. The wave-front of the transmission occupies only part of the aperture of the receiving antenna, so that the power captured by the aperture results from only that part of the wave-front that is seen by the aperture. To quantify this partial aperture area occupied by the wave-front, *effective aperture area*, A_e, is defined as the ratio of the available power at the terminals of the receiver antenna to the power flux density at the location of the receiver antenna. In general, A_e depends on the physical characteristics of the receiver antenna and the distance x between the transmitter and the receiver antennas. In References [1] and [2], A_e is assumed to be independent of x and is taken to be 4π meter^2, so that the received power expression simplifies to $P_{\rm t}/x^2$. In fact, as $x \to 0$, the power flux density approaches infinity, and with the constant A_e assumption, the received power also approaches infinity. This shows that A_e should not be taken as a constant for small values of x. In fact, A_e should approach zero as x approaches zero, so that the received power, $A_eP_{\rm t}/(4\pi x^2)$, never exceeds $P_{\rm t}$.

[||]The corresponding A_e for this model in free space is equal to $4\pi x^2/(1+x)^2$ meter^2. Note that this expression converges to 4π meter^2 as x becomes large, which is also the assumed aperture area in the conventional model.

Copyright © 2004 John Wiley & Sons, Ltd.

from different sources. Finally, we denote the average number of hops between the source and the destination of a bit by $\bar{H}$ and we let it be any real number larger than or equal to 1.

3. Definitions

We define a transmission at an arbitrary time t to be a *successful transmission*, if the SINR at the intended receiver of the transmission at time t is not smaller than β. We denote the number of simultaneously successful transmissions at time t by N_{t}. We define the *simultaneous transmission capacity of the network*, $N_{\mathrm{t}}^{\max}$, as the maximum value of N_{t} over all the placements of the N nodes, the choices of transmitters, their intended receivers and the transmission powers. Next, we define the *simultaneous transmission capacity of the network domain*, $N_{\mathrm{t}}^{\boldsymbol{Q}}$ as the maximum value of N_{t} over all the placements of the nodes, the choices of the transmitters, their intended receivers and the transmission powers, given that there are no restrictions on the number of nodes in the network. An immediate consequence of these definitions is that $N_{\mathrm{t}}^{\max} = N_{\mathrm{t}}^{\boldsymbol{Q}}$.

Let $b_i\,(T)$ be the total amount of bits of information generated by node i and received by its destinations during a T second time interval [0,T]. We define the end-to-end throughput of node i, λ_i, as follows:

$$\lambda_i := \lim_{T\to\infty} \frac{b_i(T)}{T}, \qquad 1 \le i \le N.$$

We also define the *per-node average end-to-end throughput*, λ, as the arithmetic mean of λ_i's, i.e.

$$\lambda := \frac{1}{N}\sum_{i=1}^{N} \lambda_i.$$

Then, we propose two achievability definitions: an end-to-end throughput λ_0 is said to be *achievable by all nodes*, if there exists a mobility pattern of the nodes, a traffic pattern, a spatial-temporal transmission scheduling policy and a temporal variation of transmission powers, so that $\lambda_i \ge \lambda_0$ for all $1 \le i \le N$. Likewise, an end-to-end throughput λ_0 is said to be *achievable on average*, if there exists a mobility pattern of the nodes, a traffic pattern, a spatial-temporal transmission scheduling policy and a temporal variation of transmission powers so that $\lambda \ge \lambda_0$. Note that if λ_0 is achievable by all nodes, then it is also achievable on average, and if λ_0 is *not achievable* on average, then it is not achievable by all nodes, either. Hence, we say that λ_0 is not achievable if λ_0 is not achievable on average.

Next, we propose two capacity definitions: the *per-node end-to-end throughput capacity*, λ_{e}, is the supremum of all end-to-end throughputs that are achievable by all nodes. The *per-node average end-to-end throughput capacity*, λ_{m}, is the supremum of all end-to-end throughputs that are achievable on average. An immediate consequence of these definitions is that $\lambda_{\mathrm{m}} \ge \lambda_{\mathrm{e}}$.

Finally, we define the asymptotic notations: let f and g be non-negative functions of a variable x. f is $O(g)$ with respect to x, if there exist positive real numbers x_0 and y_0, such that $0 \le f \le y_0 g$ for every $x \ge x_0$. f is $\Omega(g)$ with respect to x, if there exist positive real numbers x_1 and y_1, such that $0 \le y_1 g \le f$ for every $x \ge x_1$. f is $\Theta(g)$ with respect to x, if f is both $O(g)$ and $\Omega(g)$ with respect to x. We will omit the phrase 'with respect to x' when it can be understood from the context. Also, we will make use of the fact that f is $\Theta(g)$ with respect to x if $0 < \lim_{x\to\infty} \frac{f}{g} < \infty$.

4. Derivation of the Upper Bounds

4.1. Derivation of Upper Bounds on Simultaneous Transmission Capacity

In this subsection, we prove the following theorem that provides upper bounds on $N_{\mathrm{t}}^{\max}$ and $N_{\mathrm{t}}^{\boldsymbol{Q}}$:

Theorem 1. *For every time instant t, simultaneous transmission capacities of the network and the network domain have the following upper bounds:*

$$N_{\mathrm{t}}^{\max} \le N_{\mathrm{t}}^{\boldsymbol{Q}} \le U_\gamma, \tag{T1.1}$$

$$N_{\mathrm{t}}^{\max} \le N(1 + G/\beta), \tag{T1.2}$$

where

$$U_\gamma := \begin{cases} \frac{(\gamma-1)(\gamma-2)(1+G/\beta)d^2}{2\left(1+(\gamma-2)/(1+d)^{\gamma-1}-(\gamma-1)/(1+d)^{\gamma-2}\right)}, & \gamma \notin \{1,2\} \quad \text{(T1.3)} \\ \frac{(1+G/\beta)d}{2(1-\log(1+d)/d)}, & \gamma = 1 \quad \text{(T1.4)} \\ \frac{(1+G/\beta)d^2}{2(\log(1+d)-d/(1+d))}, & \gamma = 2 \quad \text{(T1.5)} \end{cases}$$

$$d := \frac{D}{\sqrt{2/3 - \sqrt{3}/(2\pi)}}. \tag{T1.6}$$

Copyright © 2004 John Wiley & Sons, Ltd.

Proof. Consider an arbitrary time instant t. Recall that N_{t} is the number of simultaneously successful transmissions at time t. Now, index each transmitter–receiver pair that belongs to the same transmission with a unique number between 1 and N_{t}. So, the receiver i is the intended receiver of transmitter i for every $1 \leq i \leq N_{\mathrm{t}}$. Let $\mathrm{SINR}_i(t)$ be the SINR at the receiver i at time t. Then

$$\mathrm{SINR}_i(t) = \frac{P_r^{ii}(t)}{\zeta_i(t) + \frac{1}{G}\sum_{\substack{j=1\\(j\neq i)}}^{N_{\mathrm{t}}} P_r^{ji}(t)}, \quad 1 \leq i \leq N_{\mathrm{t}}. \quad (1)$$

From the definition of a successful transmission, N_{t}, simultaneously successful transmissions can take place at time t if and only if

$$\begin{aligned}
&\mathrm{SINR}_i(t) \geq \beta, && 1 \leq i \leq N_{\mathrm{t}}\\
\Leftrightarrow\; &\frac{1}{G}\sum_{\substack{j=1\\(j\neq i)}}^{N_{\mathrm{t}}} P_r^{ji}(t) \leq \frac{P_r^{ii}(t)}{\beta} - \zeta_i(t), && 1 \leq i \leq N_{\mathrm{t}}\\
\overset{(a)}{\Leftrightarrow}\; &\frac{1}{G}\sum_{\substack{j=1\\(j\neq i)}}^{N_{\mathrm{t}}} P_{\mathrm{t}}^{j}(t)a(d_{ji}(t)) \leq \frac{P_r^{ii}(t)}{\beta} - \zeta_i(t), && 1 \leq i \leq N_{\mathrm{t}}\\
\overset{(b)}{\Leftrightarrow}\; &\frac{1}{G}\sum_{\substack{j=1\\(j\neq i)}}^{N_{\mathrm{t}}} \frac{P_{\mathrm{t}}^{j}(t)a(d_{ji}(t))}{P_{\mathrm{t}}^{i}(t)a(d_{ii}(t))} \leq \frac{1}{\beta} - \frac{\zeta_i(t)}{P_{\mathrm{t}}^{i}(t)a(d_{ii}(t))}, && 1 \leq i \leq N_{\mathrm{t}},
\end{aligned} \quad (2)$$

where step (a) follows from (a.1), and step (b) follows from dividing both sides by $P_r^{ii}(t) = P_{\mathrm{t}}^{i}(t)a(d_{ii}(t))$.

In general, for $0 \leq x \leq y + z$, $y \geq 0$ and $z \geq 0$

$$\begin{aligned}
\frac{1}{(1+x)^{\gamma}} \geq \frac{1}{(1+y+z)^{\gamma}} &\geq \frac{1}{(1+y+z+yz)^{\gamma}}\\
&= \frac{1}{(1+y)^{y}(1+z)^{\gamma}}.
\end{aligned}$$

Therefore, for $a(x)$ as defined by (a.2)

$$a(x) \geq a(y)a(z),\; 0 \leq x \leq y+z,\; y \geq 0 \quad \text{and} \quad z \geq 0. \quad (3)$$

Now, let $l_{ji}(t)$ be the distance between the receiver j and the receiver i at time t. Then, from triangle inequality

$$d_{jj}(t) \leq d_{ji}(t) + l_{ji}(t), \quad 1 \leq i \leq N_{\mathrm{t}} \text{ and } 1 \leq j \leq N_{\mathrm{t}}.$$

Setting $x = d_{ji}(t)$, $y = d_{jj}(t)$ and $z = l_{ji}(t)$ in Equation (3) we obtain that

$$a(d_{ji}(t)) \geq a(d_{jj}(t))a(l_{ji}(t)). \quad (4)$$

Thus, from Inequalities (2) and (4), if N_{t} simultaneously successful transmissions can take place at time t, then

$$\begin{aligned}
&\sum_{\substack{j=1\\(j\neq i)}}^{N_{\mathrm{t}}} \frac{P_{\mathrm{t}}^{j}(t)a(d_{jj}(t))a(l_{ji}(t))}{P_{\mathrm{t}}^{i}(t)a(d_{ii}(t))}\\
&\quad \leq \frac{G}{\beta} - \frac{G\zeta_i(t)}{P_{\mathrm{t}}^{i}(t)a(d_{ii}(t))}, \quad 1 \leq i \leq N_{\mathrm{t}}.
\end{aligned} \quad (5)$$

Define,

$$p_{ji}(t) := \frac{P_{\mathrm{t}}^{j}(t)a(d_{jj}(t))}{P_{\mathrm{t}}^{i}(t)a(d_{ii}(t))}. \quad (6)$$

Next, we add all inequalities in (5), while incorporating (6). Hence, if N_{t} simultaneously successful transmissions can take place at time t, then

$$\begin{aligned}
&\sum_{i=1}^{N_{\mathrm{t}}}\sum_{\substack{j=1\\(j\neq i)}}^{N_{\mathrm{t}}} a(l_{ji}(t))p_{ji}(t) \leq \frac{GN_{\mathrm{t}}}{\beta} - \sum_{i=1}^{N_{\mathrm{t}}} \frac{G\zeta_i(t)}{P_{\mathrm{t}}^{i}(t)a(d_{ii}(t))}\\
\Rightarrow\; &\sum_{i=1}^{N_{\mathrm{t}}-1}\sum_{j=i+1}^{N_{\mathrm{t}}} \big(a(l_{ji}(t))p_{ji}(t) + a(l_{ij}(t))p_{ij}(t)\big) \leq \frac{GN_{\mathrm{t}}}{\beta}\\
\overset{(a)}{\Leftrightarrow}\; &\sum_{i=1}^{N_{\mathrm{t}}-1}\sum_{j=i+1}^{N_{\mathrm{t}}} a(l_{ij}(t))\big(p_{ji}(t) + p_{ij}(t)\big) \leq \frac{GN_{\mathrm{t}}}{\beta}\\
\overset{(b)}{\Rightarrow}\; &\sum_{i=1}^{N_{\mathrm{t}}-1}\sum_{j=i+1}^{N_{\mathrm{t}}} 2a(l_{ij}(t)) \leq \frac{GN_{\mathrm{t}}}{\beta}\\
\overset{(c)}{\Leftrightarrow}\; &\sum_{i=1}^{N_{\mathrm{t}}-1}\sum_{j=i+1}^{N_{\mathrm{t}}} \big(a(l_{ji}(t)) + a(l_{ij}(t))\big) \leq \frac{GN_{\mathrm{t}}}{\beta}\\
\Leftrightarrow\; &\sum_{i=1}^{N_{\mathrm{t}}}\sum_{\substack{j=1\\(j\neq i)}}^{N_{\mathrm{t}}} a(l_{ji}(t)) \leq \frac{GN_{\mathrm{t}}}{\beta}\\
\overset{(d)}{\Leftrightarrow}\; &\sum_{i=1}^{N_{\mathrm{t}}}\sum_{\substack{j=1\\(j\neq i)}}^{N_{\mathrm{t}}} a(l_{ij}(t)) \leq \frac{GN_{\mathrm{t}}}{\beta},
\end{aligned} \quad (7)$$

where steps (a), (c) and (d) follow from the fact that $a(l_{ji}(t)) = a(l_{ij}(t))$ for every i, j, t and step (b) follows from the fact that $x + 1/x \geq 2$ for every positive real number x.

Copyright © 2004 John Wiley & Sons, Ltd.

From Inequality (7), we observe that the problem of obtaining an upper bound on N_t reduces to finding a lower bound on the summation term on the left-hand side of the last inequality. This term involves the sum of the attenuation function evaluated at the inter-receiver distances defined by the placement of the $N_t(N_t-1)/2$ pairs of receivers. To find the lower bound, we make use of Lemma 1, which is derived in the next subsection.

4.1.1. *Interpoint distance sum inequality*

In this subsection, we derive a lemma that gives an upper bound on the sum of the square of the nearest, the second nearest, ..., and the $(n-1)$st nearest neighbor distances from each of the n points that are arbitrarily located in a disk of diameter D.

Lemma 1. (*Interpoint distance sum inequality for a disk*) *Let $B(D)$ be a disk having diameter D. Let n points be arbitrarily placed in $B(D)$. Suppose each point is indexed by a distinct integer between* 1 *and n. Let l_{ij} be the Euclidean distance between point i and point j. Define the mth closest point to point i, a_{im}, and the Euclidean distance between point i and the mth closest point to point i, u_{im}, as follows:*

$$a_{i1} := \arg\min_{\substack{j\in\{1,2,\ldots,n\},\\ j\neq i}} \{l_{ij}\} \qquad 1\le i\le n,$$

$$a_{im} := \arg\min_{\substack{j\in\{1,2,\ldots,n\},\\ j\notin\{i\}\cup\{a_{ik}\}_{k=1}^{m-1}}} \{l_{ij}\} \quad 1\le i\le n \quad \textit{and} \quad 2\le m\le n-1,$$

$$u_{im} := l_{ia_{im}}, \quad 1\le i\le n \quad \textit{and} \quad 1\le m\le n-1.$$

Then

$$\sum_{i=1}^{n} u_{im}^2 \le \frac{mD^2}{c_2} \quad 1\le m\le n-1 \tag{L1.1}$$

where $\quad c_2 := \frac{2}{3} - \frac{\sqrt{3}}{2\pi}$.

Proof. The proof involves a spherical geometric approach, which is used to solve similar problems in Reference [14]. Let $B_i(x)$ denote the disk of diameter x, whose center is at point i. Consider the following sets of disks

$$R_m := \{B_i(u_{im}) : 1\le i\le n\}, \quad 1\le m\le n-1. \tag{L1.2}$$

Let us first consider the disks in R_1. All disks in R_1 are non-overlapping.** This can be proven by contradiction: suppose that there exist two points i and j such that $B_i(u_{i1})$ and $B_j(u_{j1})$ are overlapping. This, by the definition of overlapping, implies $(u_{i1}+u_{j1})/2 > l_{ij}$. Without loss of essential generality, suppose $u_{i1}\ge u_{j1}$. Then, $u_{i1} > l_{ij}$. However, this would contradict the definition of u_{i1}. Therefore, our original assumption, i.e. the existence of two overlapping disks in R_1, is invalid.

Let $A(X)$ denote the area of a region X. If X is a disk with diameter a, then $A(X)=\pi a^2/4$. Next, we find a lower bound on $f_{im} := A(B(D)\cap B_i(u_{im}))/A(B_i(u_{im}))$ for every $1\le i\le n$ and $1\le m\le n-1$. Pick any point S from the periphery of $B(D)$ and consider the following overlap ratio

$$f_{im}^S := A(B(D)\cap B_S(u_{im}))/A(B_S(u_{im})), \quad 1\le i\le n, 1\le m\le n-1.$$

Geometrical computation of f_{im}^S using Figure 1 leads to the following formula: $f_{im}^S = f(y)|_{y=\frac{u_{im}}{D}}$, where

$$f(y) := \frac{1}{\pi}\left(1-\frac{2}{y^2}\right)\arccos\left(\frac{y}{2}\right) + \frac{1}{y^2} - \frac{1}{\pi}\sqrt{\frac{1}{y^2}-\frac{1}{4}}.$$

Figure 2 shows the variation of $f(y)$ with y. Since $f(y)$ is a decreasing function of y and $u_{im}\le D$, $f_{im}^S \ge f(1)$ Also, $f_{im}\ge f_{im}^S$. Hence, by defining $c_2 := f(1)$, we obtain the following lower bound on f_{im} for every $1\le i\le n$ and $1\le m\le n-1$

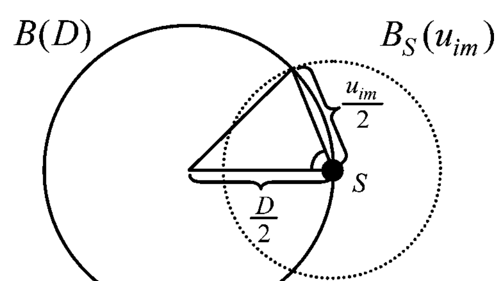


Fig. 1. Computation of the overlap ratio between $B(D)$ and $B_S(u_{im})$.

**Two disks are defined to be *overlapping* if the distance between the centers of the disks is smaller than the sum of the radii of the two disks.

Copyright © 2004 John Wiley & Sons, Ltd.

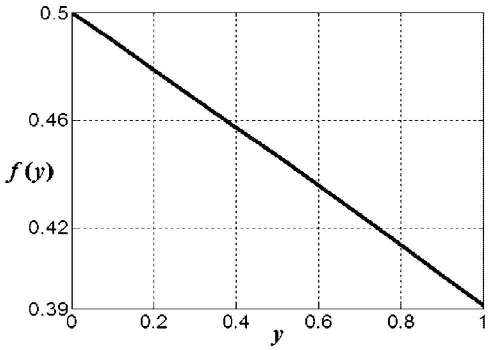


Fig. 2. Variation of the overlap ratio as a function of $y = u_{im}/D$.

$$f_{im} \geq c_2, \text{ where } c_2 = \frac{2}{3} - \frac{\sqrt{3}}{2\pi}.$$

This shows that at least c_2 fraction of the disks in R_m are inside $B(D)$. Hence, for every $1 \leq i \leq n$ and $1 \leq m \leq n-1$

$$A(B_i(u_{im}) \cap B(D)) \geq c_2 A(B_i(u_{im})) \qquad \text{(L1.3)}$$

Adding all the n inequalities in (L1.3) for a given value of m, we obtain the following inequality for every $1 \leq m \leq n-1$

$$\sum_{i=1}^{n} A(B_i(u_{im}) \cap B(D)) \geq c_2 \sum_{i=1}^{n} A(B_i(u_{im})). \qquad \text{(L1.4)}$$

Since all disks in R_1 are non-overlapping

$$\sum_{i=1}^{n} A(B_i(u_{i1}) \cap B(D)) \leq A(B(D)). \qquad \text{(L1.5)}$$

(L1.4) and (L1.5) imply

$$A(B(D)) \geq c_2 \sum_{i=1}^{n} A(B_i(u_{i1})). \qquad \text{(L1.6)}$$

In (L1.6), $A(B(D)) = \pi D^2/4$ and $A(B_i(u_{i1})) = \pi u_{i1}^2/4$. Substituting these equalities into (L1.6) and dividing both sides by $\pi c_2/4$, we obtain

$$\sum_{i=1}^{n} u_{i1}^2 \leq \frac{D^2}{c_2}. \qquad \text{(L1.7)}$$

Next, let us consider the disks in R_m for every $2 \leq m \leq n-1$. In this case, there can be overlaps between some pairs of disks in R_m. Consider two overlapping disks, $B_i(u_{im})$ and $B_j(u_{jm})$ centered at the points i and j respectively. Now, we show that if $u_{im} \geq u_{jm}$ then $j \in S_{im}$, where $S_{im} := \{a_{ik} : 1 \leq k \leq m-1\}$. This can be proven by contradiction: suppose $j \notin S_{im}$ when $u_{im} \geq u_{jm}$. Since $B_i(u_{im})$ and $B_j(u_{jm})$ are overlapping, $(u_{im} + u_{jm})/2 > l_{ij}$. Then, since $u_{im} \geq u_{jm}$, $u_{im} > l_{ij}$. However, together with our assumption $j \notin S_{im}$, this would contradict the definition of u_{im}. Therefore, our original assumption, i.e. $j \notin S_{im}$, is invalid.

Next, we show that any arbitrarily chosen point within $B(D)$ does not belong to more than m overlapping disks from R_m. The proof is again by contradiction: suppose there is a point in $B(D)$ that belongs to more than m overlapping disks from R_m. Take the largest $m+1$ of these overlapping disks. Consider the largest disk. Let b be the point at the center of this largest disk. Then all other m disks belong to S_{bm} due to the result proved in the previous paragraph. However, this contradicts the fact that the cardinality of S_{bm} is $m-1$. Therefore, our original assumption, i.e. the existence of a point belonging to more than m overlapping disks from R_m, is invalid.

Since any chosen point within $B(D)$ can belong to at most m overlapping disks from R_m, then for every $2 \leq m \leq n-1$, we have

$$\sum_{i=1}^{n} A(B_i(u_{im}) \cap B(D)) \leq mA(B(D)). \qquad \text{(L1.8)}$$

(L1.4) and (L1.8) imply

$$\sum_{i=1}^{n} u_{im}^2 \leq \frac{mD^2}{c_2}, \quad 2 \leq m \leq n-1. \qquad \text{(L1.9)}$$

Combining (L1.7) and (L1.9) completes the proof.

□

4.1.2. *Application of interpoint distance sum inequality*

In this subsection, we derive Lemma 2, which in combination with Lemma 1, provides a necessary condition for N_t simultaneously successful transmissions. Next, by using this necessary condition and another lemma, Lemma 3, we complete the derivation of the upper bound on N_t, N_t^Q and N_t^{max}.

In Lemma 1, by setting $n = N_t$ and the location of the points as the location of the receivers at time t, $u_{im}(t)$ becomes the Euclidean distance between the

Copyright © 2004 John Wiley & Sons, Ltd.

receiver i and the *mth* closest receiver to the receiver i at time t. Hence, we obtain the following inequality

$$\sum_{i=1}^{N_t}[u_{im}(t)]^2 \leq \frac{mD^2}{c_2}, \quad 1 \leq m \leq N_t - 1. \tag{8}$$

Also, from Equation (7), we obtain the following necessary condition for N_t simultaneously successful transmissions at time t

$$\begin{aligned}\sum_{m=1}^{N_t-1}\sum_{i=1}^{N_t} a(u_{im}(t)) &= \sum_{i=1}^{N_t}\sum_{\substack{j=1\\(j\neq i)}}^{N_t} a(l_{ij}(t)) \leq \frac{GN_t}{\beta}\\ \Leftrightarrow \sum_{m=1}^{N_t-1}\sum_{i=1}^{N_t}\frac{1}{(1+u_{im}(t))^{\gamma}} &\leq \frac{GN_t}{\beta}.\end{aligned} \tag{9}$$

To incorporate the constraint of Inequalities (8) into (9), we use the following lemma:

Lemma 2. *Let $x_i,\ 1 \leq i \leq n$ and C be non-negative real numbers satisfying the following inequality*

$$\sum_{i=1}^{n} x_i^2 \leq C. \tag{L2.1}$$

Let b be a non-negative real number. Then

$$\sum_{i=1}^{n}\frac{1}{(1+x_i)^b} \geq \frac{n}{\left(1+\sqrt{\frac{C}{n}}\right)^b}. \tag{L2.2}$$

Proof. For $b = 0$, (L2.2) is satisfied with equality. Thus, we consider the case when $b > 0$. Define the column vector $\boldsymbol{x}$ and the multivariate function $f(x)$ as follows:

$$\boldsymbol{x} := [x_1 x_2 \ldots x_n]^T,$$
$$f(\boldsymbol{x}) := \sum_{i=1}^{n}\frac{1}{(1+x_i)^b}.$$

We use Kuhn–Tucker Theory [15] to find the minimum value of $f(\boldsymbol{x})$ subject to the constraint in (L2.1). We define the constraint function, $g(\boldsymbol{x})$ as follows:

$$g(\boldsymbol{x}) := C - \sum_{i=1}^{n} x_i^2 \geq 0.$$

Let $\boldsymbol{y} = [y_1 y_2 \ldots y_n]^T$ be the column vector at which f takes its minimum value. Then, from Kuhn–Tucker Theory, there exists $\theta \geq 0$, such that the following conditions are satisfied

$$(\nabla f(\boldsymbol{x}) - \theta\nabla g(\boldsymbol{x}))|_{\boldsymbol{x}=\boldsymbol{y}} \geq \mathbf{0}, \tag{L2.3}$$

$$\boldsymbol{y}^T(\nabla f(\boldsymbol{x}) - \theta\nabla g(\boldsymbol{x}))|_{\boldsymbol{x}=\boldsymbol{y}} = 0, \tag{L2.4}$$

$$\theta g(\boldsymbol{y}) = 0, \tag{L2.5}$$

$$g(\boldsymbol{y}) \geq 0. \tag{L2.6}$$

From Equations (L2.3) and (L2.4) we obtain

$$\frac{-b}{(1+y_i)^{b+1}} + 2\theta y_i \geq 0 \quad 1 \leq i \leq n, \tag{L2.7}$$

$$\sum_{i=1}^{n} y_i\left(\frac{-b}{(1+y_i)^{b+1}} + 2\theta y_i\right) = 0. \tag{L2.8}$$

Since $\boldsymbol{y} \geq \mathbf{0}$, we need to determine whether the constraint is binding or not. Namely, we need to determine whether or not there exist any components of $\boldsymbol{y}$ that are zero. To do this, we compare the values of f in all possible cases. Let y has k zero components, i.e. $k := |\{i : \boldsymbol{y}_i = 0, 1 \leq i \leq n\}|$. So, $0 \leq k \leq n$. If $0 \leq k < n$ then from (L2.7) $\theta > 0$, from (L2.5) $g(\boldsymbol{y}) = 0$ and from (L2.8) all non-zero components of $\boldsymbol{y}$ are equal to each other. Thus, since $g(\boldsymbol{y}) = 0$, all non-zero components of $\boldsymbol{y}$ are equal to $[C/(n-k)]^{1/2}$. If $k = n$ then all components of $\boldsymbol{y}$ are zero and $f(\boldsymbol{y}) = n$. Hence,

$$f(\boldsymbol{y}) = \begin{cases} k + \frac{n-k}{\left(1+\sqrt{\frac{C}{n-k}}\right)^b}, & 0 \leq k < n\\ n, & k = n.\end{cases} \tag{L2.9}$$

Next, we show that the expression in (L2.9) is minimized for $k = 0$ and therefore, $\boldsymbol{y}$ has non-zero components. To show this, we prove the validity of the following inequality

$$x + \frac{n-x}{\left(1+\sqrt{\frac{C}{n-x}}\right)^b} \geq \frac{n}{\left(1+\sqrt{\frac{C}{n}}\right)^b}, \quad x \in [0, n). \tag{L2.10}$$

Define the function on the left-hand side of (L2.10) as $h(x)$. Taking the partial derivative of $h(x)$ with respect to x and using Bernoulli's Inequality [16], we obtain

$$\frac{\partial h}{\partial x} = 1 - \frac{1 + \left(\frac{b}{2}+1\right)\sqrt{\frac{C}{n-x}}}{\left(1+\sqrt{\frac{C}{n-x}}\right)^{b+1}} > 1 - \frac{1 + \left(\frac{b}{2}+1\right)\sqrt{\frac{C}{n-x}}}{1 + (b+1)\sqrt{\frac{C}{n-x}}} = \frac{\frac{b}{2}\sqrt{\frac{C}{n-x}}}{1 + (b+1)\sqrt{\frac{C}{n-x}}} \geq 0$$

So, $h(x)$ is an increasing function of x over $[0, n)$. Hence, $h(x) \geq h(0)$ for every $x \in [0, n)$. This completes the proof of (L2.10) and also implies that $f(\boldsymbol{y})$ assumes minimum value for $k = 0$. This further implies that all components of y are equal to $(C/n)^{1/2}$ and that $f(\boldsymbol{y}) = n/[1 + (C/n)^{1/2}]^b$. Since $f(\boldsymbol{x}) \geq f(\boldsymbol{y})$ for every $\boldsymbol{x} \geq \boldsymbol{0}$, (L2.2) follows. □

Next, in Lemma 2, we set $n = N_t$, $b = \gamma$, $x_i = u_{im}(t)$, $C = mD^2/c_2$ for every $1 \leq i \leq n$ and $1 \leq m \leq n-1$, so that (L2.1) and (8) become identical. Also the left-hand side of (L2.2) and the inner summation in Inequality (9) become identical. Hence, we obtain the following lower bound on the left-hand side of Inequality (9):

$$\sum_{m=1}^{N_t-1}\sum_{i=1}^{N_t}\frac{1}{(1+u_{im}(t))^{\gamma}} \geq \sum_{m=1}^{N_t-1}\frac{N_t}{\left(1+\frac{D}{\sqrt{c_2}}\sqrt{\frac{m}{N_t}}\right)^{\gamma}}. \quad (10)$$

Next, we define $d := D/c_2^{1/2}$. The quantity d is the diameter of the network domain divided by a constant approximately equal to 0.625. Combining this definition with Inequalities (9) and (10), we obtain the following necessary condition for N_t simultaneously successful transmissions at time t

$$\sum_{m=1}^{N_t-1}\frac{1}{\left(1+d\sqrt{\frac{m}{N_t}}\right)^{\gamma}} \leq \frac{G}{\beta} \overset{(a)}{\Rightarrow} \int_1^{N_t}\left(1+d\sqrt{\frac{x}{N_t}}\right)^{-\gamma}\mathrm{d}x \leq \frac{G}{\beta}$$
$$\overset{(b)}{\Leftrightarrow} \frac{2N_t}{d^2}\int_{1+\frac{d}{\sqrt{N_t}}}^{1+d}\frac{u-1}{u^{\gamma}}\mathrm{d}u \leq \frac{G}{\beta}, \quad (11)$$

where step (a) follows from the fact that $\int_a^{b+1} f(x)\mathrm{d}x \leq \sum_{m=a}^{b} f(m)$, whenever a and b are integers and $f(x)$ is a continuous and non-increasing function of x over $[a, b+1]$ and step (b) follows from changing the variable of the integration by defining $u = 1 + d(x/N_t)^{1/2}$. Next, we define σ as follows

$$\sigma := \sqrt{N_t}/d. \quad (12)$$

The quantity σ^2 is proportional to the average number of successful transmissions per unit area. Combining Inequalities (11) and (12), the necessary condition for N_t simultaneously successful transmissions at time t becomes

$$2\sigma^2\int_{1+\frac{1}{\sigma}}^{1+d}(u^{1-\gamma} - u^{-\gamma})\mathrm{d}u \leq \frac{G}{\beta}, \quad (13)$$

Next, we use the following lemma to obtain a closed form solution for the upper bound on N_t.

Lemma 3. *Let $x > 0$ and $y > 0$ be real numbers such that $xy \geq 1$ and let $a \leq 1$ be a real number. Let $I := 2x^2\int_{1+1/x}^{1+y}(u^a - u^{a-1})\mathrm{d}u$. Then,*

$$I \geq \frac{2x^2[a(1+y)^{a+1} - (a+1)(1+y)^a]}{a(a+1)} + \frac{2x^2}{a(a+1)} - 1, \quad a \leq 1 \text{ and } a \notin \{-1, 0\} \quad \text{(L3.1)}$$

$$I \geq 2x^2 y - 2x^2\log(1+y) - 1, \quad a = 0 \quad \text{(L3.2)}$$

$$I \geq 2x^2\log(1+y) - \frac{2x^2 y}{1+y} - 1, \quad a = -1. \quad \text{(L3.3)}$$

Proof. Firstly, consider the case when $a \leq 1$ and $a \notin \{1, 0\}$. For $a \in (-\infty, -1) \cup (0, 1)$, define

$$R_1(x,a) := \frac{2x^2}{2x(x-a)(1+\frac{1}{x})^a + a(a+1)}$$

and for $a \in (-1, 0)$, define

$$R_2(x,a) := \frac{2x(x-a)(1+\frac{1}{x})^a}{2x^2 - a(a+1)}.$$

R_1 and R_2 are differentiable functions of x and partial derivatives of them with respect to x are

$$\frac{\partial R_1}{\partial x} = \frac{4a(a+1)x^{2a+1}\left[(1+x)^{1-a} - x^{1-a}\right]}{(1+x)^{1-a}[2x^2(1+x)^a - 2ax(1+x)^a + a(a+1)x^a]^2} \geq 0,$$

$$\frac{\partial R_2}{\partial x} = \frac{-2a(1-a^2)(2x-a)}{(1+x)^{1-a}x^a[2x^2 - a(a+1)]^2} \geq 0.$$

Copyright © 2004 John Wiley & Sons, Ltd.

Also, $\lim_{x\downarrow 0} R_1 = \lim_{x\downarrow 0} R_2 = 0$. So, R_1 and R_2 are non-decreasing and non-negative functions of x. Finally, $\lim_{x\to\infty} R_1 = \lim_{x\to\infty} R_2 = 1$. Therefore, for $a \in (-\infty, -1) \cup (0, 1)$,

$$\begin{aligned} R_1 &= \frac{2x^2}{2x(x-a)(1+\frac{1}{x})^a + a(a+1)} \leq 1 \\ & a \in (-\infty, -1) \cup (0, 1] \\ &\Rightarrow \frac{2x(x-a)(1+\frac{1}{x})^a}{a(a+1)} \geq \frac{2x^2}{a(a+1)} - 1 \end{aligned} \tag{L3.4}$$

and for $a \in (-1, 0)$,

$$\begin{aligned} R_2 &= \frac{2x(x-a)(1+\frac{1}{x})^a}{2x^2 - a(a+1)} \leq 1, \\ &\Rightarrow \frac{2x(x-a)(1+\frac{1}{x})^a}{a(a+1)} \geq \frac{2x^2}{a(a+1)} - 1. \end{aligned} \tag{L3.5}$$

Combining (L3.4) and (L3.5) we find that for $a \leq 1$ and $a \notin \{-1, 0\}$,

$$\begin{aligned} &\frac{2x^2}{a(a+1)} - 1 \leq \frac{2x(x-a)(1+\frac{1}{x})^a}{a(a+1)} \\ &\Rightarrow \frac{2x^2[a(1+y)^{a+1} - (a+1)(1+y)^a]}{a(a+1)} \\ &\quad + \frac{2x^2}{a(a+1)} - 1 \\ &\leq \frac{2x^2[a(1+y)^{a+1} - (a+1)(1+y)^a]}{a(a+1)} \\ &\quad + \frac{2x(x-a)(1+\frac{1}{x})^a}{a(a+1)} \\ &= 2x^2 \left\{ \frac{1}{(a+1)} \left[(1+y)^{a+1} - \left(1+\frac{1}{x}\right)^{a+1} \right] \right. \\ &\quad \left. - \frac{1}{a} \left[(1+y)^a - \left(1+\frac{1}{x}\right)^a \right] \right\} \\ &= I. \end{aligned}$$

This completes the proof of (L3.1). We prove (L3.2) and (L3.3) as follows: Let $i \in \{-1, 0\}$. Then

$$\begin{aligned} I|_{a=i} &\stackrel{(a)}{=} \lim_{a\uparrow i} I \\ &\stackrel{(b)}{\geq} \lim_{a\uparrow i} \left\{ \frac{2x^2[a(1+y)^{a+1} - (a+1)(1+y)^a]}{a(a+1)} \right. \\ &\quad \left. + \frac{2x^2}{a(a+1)} - 1 \right\} \\ &\stackrel{(c)}{=} \begin{cases} 2x^2 y - 2x^2 \log(1+y) - 1, & i = 0, \\ 2x^2 \log(1+y) - \frac{2x^2 y}{1+y} - 1 & i = -1 \end{cases} \end{aligned}$$

where step (a) follows from the continuity of I at $a = i$, step (b) follows from (L3.1) and step (c) follows from L'Hôpital's Rule [17]. This completes the proof of Lemma 3. □

By setting $a = 1 - \gamma, x = \sigma$ and $y = d$ in Lemma 3, I becomes equal to the left-hand side of Equation (13). Hence, using (L3.1), (L3.2) and (L3.3), we obtain the following necessary conditions for N_{t} simultaneously successful transmissions at time t

$$\begin{aligned} &\frac{2\sigma^2\left(\frac{1-\gamma}{(1+d)^{\gamma-2}} - \frac{2-\gamma}{(1+d)^{\gamma-1}}\right)}{(1-\gamma)(2-\gamma)} + \frac{2\sigma^2}{(1-\gamma)(2-\gamma)} - 1 \\ &\quad \leq \frac{G}{\beta}, \quad \gamma \notin \{1, 2\} \end{aligned} \tag{14}$$

$$2\sigma^2 d - 2\sigma^2 \log(1+d) - 1 \leq \frac{G}{\beta}, \quad \gamma = 1 \tag{15}$$

$$2\sigma^2 \log(1+d) - \frac{2\sigma^2 d}{1+d} - 1 \leq \frac{G}{\beta}, \quad \gamma = 2. \tag{16}$$

Solving Inequalities (14), (15) and (16) for σ and substituting σ from Inequality (12), we obtain the following upper bound on N_{t}

$$N_{\mathrm{t}} \leq U_\gamma, \tag{17}$$

where U_γ is defined as in (T1.3), (T1.4) and (T1.5).

Recall that $N_{\mathrm{t}}^{\boldsymbol{Q}}$ is the maximum value of N_{t} over all the placements of the nodes, the choice of the transmitters, their intended receivers and the transmission powers, given that there are no restrictions on the value of N. Since there have been no restrictions on any of these parameters during the derivation of Equation (17), hence, the right-hand side of Equation (17) is also an upper bound on $N_{\mathrm{t}}^{\boldsymbol{Q}}$, which is not less than $N_{\mathrm{t}}^{\max}$. This completes the proof of (T1.1).

Finally, we complete the proof of (T1.2) as follows: suppose there is a single receiver node and N_{t} transmissions intended for this node at time t. Then, in Inequality (7), $l_{ij}(t)$ is equal to zero for every i, j and t. Thus, N_{t} will be not more than $1+G/\beta$. This shows that none of the N nodes can receive more than $1+G/\beta$ simultaneously successful transmissions intended for itself and, thus, $N_{\mathrm{t}}^{\max}$ cannot exceed $N(1+G/\beta)$. This completes the proof of Theorem 1. □

Copyright © 2004 John Wiley & Sons, Ltd.

4.2. An Upper Bound on Simultaneous Transmission Capacity Implies an Upper Bound on Per-Node End-to-End Throughput Capacity

In this subsection, firstly, we prove the following theorem:

Theorem 2. λ_e *and* λ_m *are upper bounded as follows:*

$$\lambda_e \le \lambda_m \le \frac{W_{\max} U_\gamma}{\bar{H} N}, \tag{T2.1}$$

$$\lambda_e \le \lambda_m \le \frac{W_{\max}}{\bar{H}}\left(1 + \frac{G}{\beta}\right). \tag{T2.2}$$

Proof. Define the *total information transmission rate of the network at time t*, $C(t)$, as follows:

$$C(t) := \sum_{i=1}^{N_t} W_i(t),$$

where $W_i(t)$ is the transmission rate of the i^{th} successful transmission at time t. By the definition of $\bar{H}$ in Section 2.4, each bit of information delivered to its destination is transmitted in $\bar{H}$ hops on the average. Therefore, the time average of $C(t)$ over $t \in [0, \infty)$ is not less than $\bar{H}\sum_{i=1}^{N} \lambda_i = \bar{H}N\lambda$. Also, since $W_i(t) \le W_{\max}$ and $N_t \le N_t^{\max}$, $C(t) \le W_{\max} N_t^{\max}$. Thus, average of $C(t)$ over any time interval cannot exceed $W_{\max}$ multiplied by a time invariant upper bound on $N_t^{\max}$, such as the upper bounds in (T1.1) and (T1.2). Hence

$$\bar{H}N\lambda \le W_{\max} U,$$

where $U \in \{U_\gamma, N(1 + G/\beta)\}$. As a result, we obtain the following upper bound on λ:

$$\lambda \le \frac{W_{\max} U}{\bar{H} N}. \tag{18}$$

It is worth emphasizing that, because of the generality of the network model underlying the derivation of Inequality (18), it is applicable even when the mobility pattern of the nodes, the spatial–temporal transmission scheduling policy, the temporal variation of transmission powers, the source–destination pairs and the possibly multi-path routes between them are optimally chosen as to maximize λ. Similarly, Inequality (18) is applicable even when the nodes are capable of maintaining multiple transmissions and/or receptions simultaneously.

Recall that λ_m is the supremum of all end-to-end throughputs λ_0 for which there exist: a mobility pattern of the nodes, a traffic pattern, a spatial-temporal transmission scheduling policy and a temporal variation of transmission powers, so that $\lambda \ge \lambda_0$. There have been no restrictions on these parameters during the derivation of Inequality (18), hence the right-hand side of Inequality (18) is also an upper bound on λ_m, which is not less than λ_e. This completes the proof of Theorem 2. □

So far, there have been no restrictions on the number of simultaneous transmissions and/or receptions that a node is capable of maintaining. If, as in References [1] and [2], there is also the additional restriction that a node cannot transmit and receive simultaneously and that a node is capable of maintaining at most one transmission or one reception at any given time, then $N_t^{\max} \le N/2$ and thus, $\lambda \le W_{\max}/(2\bar{H})$. We will henceforth refer to this latter case as the *half-duplex restricted case*. Combining this condition with Inequality (18) leads to the following upper bound on λ for the half-duplex restricted case

$$\lambda \le \frac{W_{\max}}{\bar{H}} \min\left\{\frac{U_\gamma}{N}, \frac{1}{2}\right\}. \tag{19}$$

Hence, proceeding in a similar way as in the proofs of (T2.1) and (T2.2), we conclude that the right-hand side of Inequalitiy (19) is also an upper bound on λ_e and λ_m for the half-duplex restricted case.

Finally, we show that dividing the communication bandwidth into several sub-channels of smaller bandwidth does not change the terms other than $W_{\max}$ in all of the results that we have presented so far. An assumption behind the results is that all transmissions are taking place in the same communication bandwidth. If the communication bandwidth is divided into several sub-channels of smaller bandwidth, then there still is an upper bound on the transmission rate in each of these sub-channels. All of the upper bounds on simultaneous transmission capacities of $\boldsymbol{Q}$ and the network are still valid for each of these sub-channels individually. Therefore, if there are M sub-channels and the transmission rate of the m^{th} sub-channel is not more than $W_m^{\max}$, then all of the upper bounds on λ_e and λ_m are still valid if $W_{\max}$ is replaced with $\sum_{m=1}^{M} W_m^{\max}$.

Copyright © 2004 John Wiley & Sons, Ltd.

5. Analysis of the Upper Bounds

In this section, firstly, we analyze the asymptotic and limiting behavior of the upper bound U_γ in Theorem 1, to draw the following conclusions about $N_t^{\boldsymbol{Q}}$

- $\lim\limits_{\gamma\to\infty} \frac{U_\gamma}{\gamma^2} = \frac{1}{2c_2}(1+\frac{G}{\beta})D^2 \Rightarrow N_t^{\boldsymbol{Q}}$ is $O(\gamma^2)$
- $\lim\limits_{D\to\infty} \frac{U_\gamma|_{\gamma=1}}{D} = \frac{1}{2{c_2}^{1/2}}(1+\frac{G}{\beta}) \Rightarrow N_t^{\boldsymbol{Q}}$ is $O(D)$ if $\gamma=1$
- $\lim\limits_{\substack{D\to\infty \\ (\gamma<2)}} \frac{U_\gamma}{D^\gamma} = \frac{1}{c_2^{\gamma/2}}(1-\frac{\gamma}{2})(1+\frac{G}{\beta}) \Rightarrow N_t^{\boldsymbol{Q}}$ is $O(D^\gamma)$ if $\gamma<2$
- $\lim\limits_{D\to\infty} \frac{U_\gamma|_{\gamma=2}}{D^2/\log(D)} = \frac{1}{2c_2}(1+\frac{G}{\beta}) \Rightarrow N_t^{\boldsymbol{Q}}$ is $O(D^2/\log(D))$ if $\gamma=2$
- $\lim\limits_{\substack{D\to\infty \\ (\gamma>2)}} \frac{U_\gamma}{D^2} = \frac{(\gamma-1)(\gamma-2)}{2c_2}\left(1+\frac{G}{\beta}\right) \Rightarrow N_t^{\boldsymbol{Q}}$ is $O(D^2)$ if $\gamma>2$
- $\lim\limits_{G/\beta\to\infty} \frac{U_\gamma}{G/\beta} = f(\gamma,d) := \frac{U_\gamma}{1+G/\beta} \Rightarrow N_t^{\boldsymbol{Q}}$ is $O(G/\beta)$
- $\lim\limits_{\gamma\downarrow 0} U_\gamma = \lim\limits_{D\downarrow 0} U_\gamma = 1+\frac{G}{\beta} \Rightarrow$ Lack of attenuation is equivalent to lack of space.

Also, since the area of the network domain is $A=\pi D^2/4$, D can be replaced with $(4A/\pi)^{1/2}$. Doing so, we can also conclude that $N_t^{\boldsymbol{Q}}$ is $O(A^{\min\{\gamma/2,1\}})$ if $\gamma\neq 2$ and $O(A/\log(A))$ if $\gamma=2$. Regardless of the value of γ, this also implies that $N_t^{\boldsymbol{Q}}$ cannot grow with the area of $\boldsymbol{Q}$ super-linearly. Linear growth is not possible when $\gamma\leq 2$ and can only be possible when $\gamma>2$.

In Figure 3, U_γ is plotted as a function of A and γ, for $G=\beta=10$. This figure illustrates the growth trend of U_γ as γ and/or A increase. It is possible to observe the linear and the sub-linear growth of U_γ with A when $\gamma>2$ and $0<\gamma\leq 2$ respectively. The figure also illustrates the equivalence of the lack of attenuation ($\gamma=0$) and the lack of space ($A=0$). One should also notice the quadratic growth of U_γ with γ.

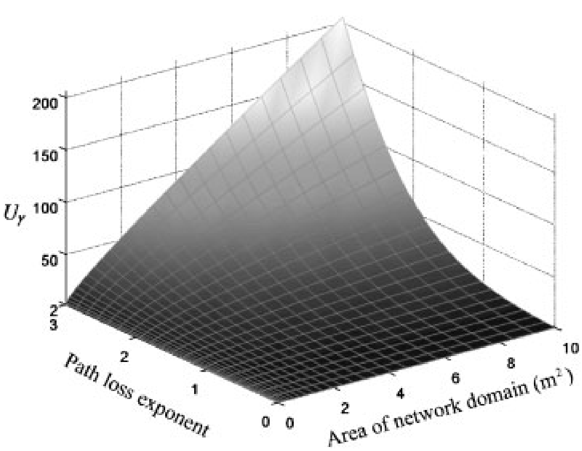


Fig. 3. Upper bound on the simultaneous transmission capacity of the network domain as a function of area of the network domain and the path loss exponent.

Secondly, we analyze the upper bounds on $N_t^{\max}$. (T1.1) of Theorem 1 shows that $N_t^{\max}$ is $O(1)$ with respect to N. Since $N_t^{\max}\leq N_t^{\boldsymbol{Q}}$, all of the above asymptotic results are valid for $N_t^{\max}$, too.

However, from (T1.2), $N_t^{\max}\leq N(1+G/\beta)$. Therefore, for a given N, G, and β, the upper bound on $N_t^{\max}$ in (T1.1) loses its tightness beyond some finite values of D and γ. Existence of an upper bound on $N_t^{\max}$ independent of D and γ also shows that $N_t^{\max}$ is $O(1)$ with respect to A and γ. The reason is that beyond some finite values of A or γ, the network domain provides sufficient space and attenuation, so that the upper bound on the number of simultaneous receptions per-node, i.e. $1+G/\beta$, becomes the limiting factor.

Next, we analyze asymptotic and limiting behavior of the upper bounds on λ_e and λ_m. (T2.1) of Theorem 2 shows that λ_e and λ_m are $O(1/N)$ and $O(1/\bar{H})$. It also shows[††] that λ_e and λ_m are $O(G/\beta)$. We also observe that λ_e and λ_m are upper bounded by $W_{\max}(1+G/\beta)/(\bar{H}N)$ when the network domain lacks attenuation or space. Due to (T2.2), λ_e and λ_m cannot exceed $W_{\max}(1+G/\beta)/\bar{H}$, which is independent of A and γ. So, $W_{\max}(1+G/\beta)/\bar{H}$ becomes the dominant upper bound beyond some finite values of A or γ and thus, λ_e and λ_m are $O(1)$ with respect to A and γ. Similar behavior is also observable in the half-duplex restricted case; beyond some finite values of A or γ, the network domain provides sufficient space and attenuation so that $\lfloor N/2\rfloor$ simultaneously successful transmissions become possible.[‡‡] However, no more transmission can be scheduled, since there are no remaining inactive pairs of nodes, and thus, λ_e and λ_m cannot exceed $W_{\max}/(2\bar{H})$. In general, there is a region of (A,γ) pairs for which the dominant upper bound on λ_e and λ_m is $W_{\max}U_\gamma/(\bar{H}N)$. This region is bounded by the A axis, the γ axis and the set of (A,γ) pairs for which $U_\gamma/N=1+G/\beta$. Since U_γ is an increasing function of A and γ, this region will expand as N increases. This shows that the limitation of λ_e and λ_m due to shortage of space and attenuation is more pronounced when N is large compared to U_γ. Additionally, we have shown that U_γ is $\Theta(A^{\min\{\gamma/2,1\}})$ when

[††] The $O(G/\beta)$ result assumes that $W_{\max}$ is not dependent on G/β. However, in some practical systems, $W_{\max}$ is inversely proportional to G/β, as we will see in Section 7.

[‡‡] $\lfloor x\rfloor$ denotes the largest integer smaller than or equal to x.

Copyright © 2004 John Wiley & Sons, Ltd.

$\gamma \neq 2$, $\Theta(A/\log(A))$ when $\gamma = 2$ and also $\Theta(\gamma^2)$. These observations support the claim that for large N, there is a region of (A,γ) pairs where additional space and additional attenuation provide considerable increase in λ_e and λ_m, where the behavior of λ_e and λ_m resembles the asymptotic behavior of U_γ and beyond this region the behavior λ_e and λ_m changes into $\Theta(1)$ with respect to A and γ.

Next, we demonstrate the above results through an example. Consider the half-duplex restricted case. We have shown that λ_e and λ_m cannot exceed the right-hand side of Inequality (19). Now, we normalize this quantity with respect to W_{max} and we denote the resulting expression by Λ_U. In Figure 4, Λ_U is plotted as a function of A and γ. The other parameters for this example are: $G = \beta = 10$, $N = 250$ and $\bar{H} = 1$.[§§] This figure illustrates the variation in the growth trend of Λ_U as a function of A for various values of γ. Also, it demonstrates the presence of a region of (A,γ) pairs where the limitation of λ_e and λ_m is due to shortage of space and attenuation. For the (A,γ) pairs outside of this region, shortage of inactive pairs of nodes becomes the dominant limitation and thus, $W_{max}/(2\bar{H})$ becomes the dominant upper bound on λ_e and λ_m.

In Figure 5, parameter values are the same except that N is now an independent variable and $\gamma = 3$. The green region consists of the (A,N) pairs where the limitation of λ_e and λ_m is due to shortage of space. For the (A,N) pairs outside of this region, namely inside the blue region, shortage of inactive pairs of nodes is the dominant limitation, and thus, $W_{max}/(2\bar{H})$ is the dominant upper bound on λ_e and λ_m. The figure also demonstrates that if the area of the network domain is kept constant and the number of nodes is increased, then Λ_U decays as $\Theta(1/N)$, so that λ_e and λ_m vanish as N grows large. However, if the area also increases with N, we observe that it can be possible to keep Λ_U at a constant level so that it does not rule out the possibility of achieving a non-vanishing per-node end-to-end throughput as the number of nodes grows large. We will elaborate on this result in Section 7.

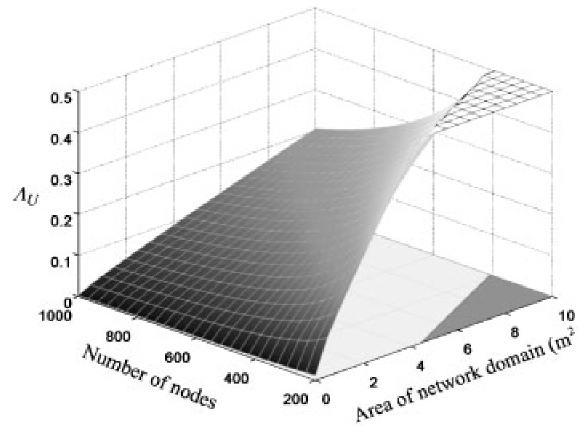


Fig. 5. Upper bound on the normalized per-node end-to-end throughput capacity as a function of area of the network domain and number of nodes in the network.

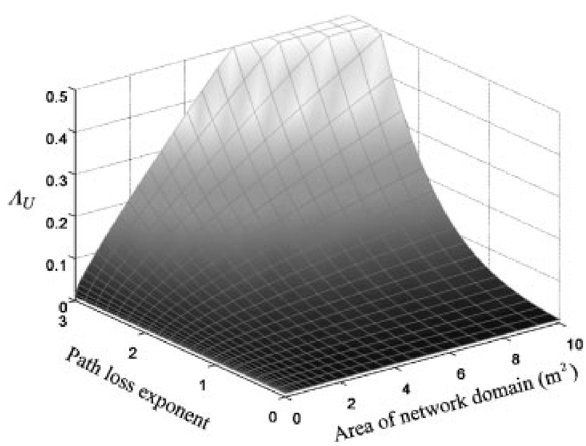


Fig. 4. Upper bound on the normalized per-node end-to-end throughput capacity as a function of area of the network domain and the path loss exponent.

6. λ_e and λ_m are $\Theta(1/N)$

In the previous section, we have shown that λ_e and λ_m are $O(1/N)$. Next, to prove they are also $\Theta(1/N)$, we show that they are $\Omega(1/N)$. We do this by constructing a TDMA scheme that assigns each of the nodes a separate time slot of constant duration. In such a scheme, there are N slots in each cycle and each of the nodes transmits directly to its destination in the slot assigned to itself, with a transmission power large enough to satisfy the signal to noise ratio requirement. Assuming ζ is an upper bound on the power of noise in the used communication bandwidth, a transmission power of $\beta\zeta/a(D)$ guarantees successful reception.

Although this simple scheme takes no advantage of spatial reuse, it allows each of the nodes transmitting $1/N$ fraction of the time. Thus, assuming that each transmission satisfying the signal to noise ratio requirement can occur with rate W, a per-node

[§§]This is the least possible value $\bar{H}$ can take and achieved when each bit of generated information is destined for a node one hop away.

Copyright © 2004 John Wiley & Sons, Ltd.

end-to-end throughput of W/N is achievable by all nodes. This shows that λ_e and λ_m are $\Omega(1/N)$. As a result, λ_e and λ_m are $\Theta(1/N)$.

7. Implications of the Results on Scalability

In this section, we consider the following scalability problem: we are increasing the number of nodes in the network indefinitely and we want to achieve a desired per-node end-to-end throughput, say λ_0. λ_0 is not achievable if no other parameter is increased as a function of N, since λ_e and λ_m are no more than $W_{\max}U_\gamma/(\bar{H}N)$, which is $O(1/N)$. So, one or more of the parameters from $W_{\max}$, γ, G/β or A must increase with N and N must be increasing according to a function of U_γ, so that $W_{\max}U_\gamma/(\bar{H}N) \geq \lambda_0$. Note that $\bar{H}$ cannot be indefinitely reduced to compensate for increasing N, because $\bar{H} \geq 1$, as every bit of information has to be transmitted for at least one hop. This shows that $\bar{H}N$ must be $O(W_{\max}U_\gamma)$.

For practical systems, γ is a property of the wireless channel and it cannot increase with N. $W_{\max}$ cannot increase indefinitely with N, because of the presence of noise and because of the maximum transmission power constraints. These limit reliable information transmission to rates that do not grow with N. On the other hand, G/β depends on the implementation of the communication system and increasing it for a given system bandwidth usually requires decreasing $W_{\max}$. For example, it is shown in Reference [10] that in spread spectrum CDMA, for a given system bandwidth, symbol transmission rate is inversely proportional to the processing gain. Likewise, reducing β requires a proportional decrease in the symbol transmission rate to satisfy a given bit error rate requirement. Therefore, increasing G/β will not compensate for increasing N. So, the only way of achieving λ_0 would be increasing A as N increases. Hence, N must be increasing as a function of A. We have shown that U_γ is $\Theta(A^{\min\{\gamma/2,1\}})$ when $\gamma \neq 2$ and $\Theta(A/\log(A))$ when $\gamma = 2$. Therefore, unless N is $O(A^{\min\{\gamma/2,1\}})$ when $\gamma \neq 2$ and $O(A/\log(A))$ when $\gamma = 2$, λ_0 is not achievable. Also, $\bar{H}$ must be $\Theta(1)$ with respect to N due to the following reasoning. We know that $\bar{H} \geq 1$, which implies that $\bar{H}$ is $\Omega(1)$. To see why $\bar{H}$ must be $O(1)$ with respect to N, recall that λ_0 cannot exceed $W_{\max}(1+G/\beta)/\bar{H}$ and increasing G/β requires a proportional reduction in $W_{\max}$, as is the case in spread spectrum CDMA. Thus, compensating for indefinitely growing $\bar{H}$ by increasing G/β is not possible.

The above results can also be stated in terms of *node density*, $\rho := N/A$. From the above paragraph, dividing N and the asymptotic upper bounds on N by A, we obtain the following result: unless ρ is $O(A^{\min\{\gamma/2-1,0\}})$ when $\gamma \neq 2$ and $O(1/\log(A))$ when $\gamma = 2$, λ_0 is not achievable. In other words, λ_0 is not achievable if ρ grows with N indefinitely when $\gamma > 2$, if $\rho\log(A)$ grows with N indefinitely when $\gamma = 2$, and if $\rho A^{1-\gamma/2}$ grows with N indefinitely when $\gamma < 2$. In any case, λ_0 is not achievable if ρ grows with N indefinitely. Also, when $\gamma \leq 2$, unless ρ decays down to 0 as $N \to \infty$, λ_0 is not achievable. Our observations in this and the previous paragraphs prove the following corollary regarding practical systems.

Corollary. (*Necessary condition for scalability of practical systems*) *A desired per-node end-to-end throughput is not achievable as* $N \to \infty$, *unless H is* $\Theta(1)$ *with respect to N, A also grows with N and the following equivalent conditions are satisfied.*

- *N is* $O(A^{\min\{\gamma/2,1\}})$ *when* $\gamma \neq 2$ *and* $O(A/\log(A))$ *when* $\gamma = 2$, *or, equivalently*
- ρ *is* $O(A^{\min\{\gamma/2-1,0\}})$ *when* $\gamma \neq 2$ *and* $O(1/\log(A))$ *when* $\gamma = 2$.

Figure 6 illustrates this corollary. In this figure, $G = \beta = 10$, $\bar{H} = 1$ and the curves are obtained by plotting the (A,N) pairs, for which $\Lambda_U = 0.1$ and $\gamma \in \{0,1,2,3\}$. We know that normalized λ_e and λ_m are no more than Λ_U, which is a decreasing function of N and an increasing function of A when $\Lambda_U\bar{H} < 0.5$. Therefore, each of these curves separates a region of (A,N)

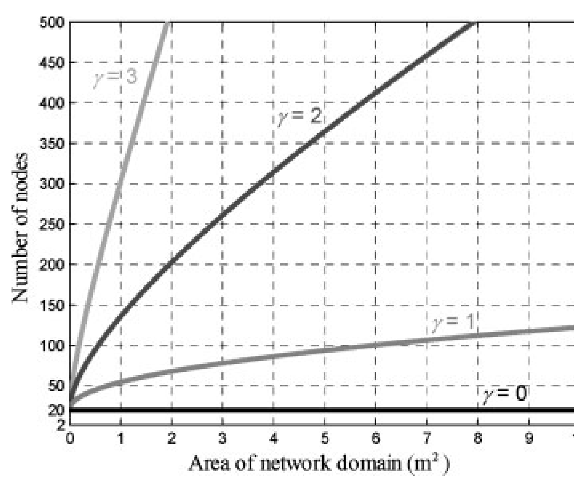


Fig. 6. Curves formed by the (A,N) pairs for which $\Lambda_U = 0.1$. For the (A,N) pairs above the curves, any normalized throughput greater than or equal to 0.1 is not achievable.

Copyright © 2004 John Wiley & Sons, Ltd.

pairs where a normalized end-to-end throughput of 0.1 is not achievable and another region where it may be achievable on average or by all nodes. For example, when $\gamma = 2$ and $(A,N) = (3{,}400)$, the normalized end-to-end throughput 0.1 is not achievable, whereas it may be achievable on average or by all nodes for $(A,N) = (3{,}100)$. The corollary tells us that for the sequence of (A,N) pairs forming each of the curves in Figure 6, N is $\Theta(1)$, $\Theta(A^{1/2})$, $\Theta(A/\log(A))$ and $\Theta(A)$ when γ is 0, 1, 2 and 3 respectively. Equivalently, for the sequence of (A,ρ) pairs associated with each of these curves, ρ is $\Theta(1/A)$, $\Theta(1/A^{1/2})$, $\Theta(1/\log(A))$ and $\Theta(1)$ when γ is 0, 1, 2 and 3 respectively.

8. Conclusions

In this paper, we have studied the capacity of wireless networks with a more general network model than the models used in References [1] and [2], and we have presented the implications of our results on network scalability.

Instead of the propagation model used in References [1] and [2], we used the power law decaying propagation model, which was proposed in other studies such as References [12] and [13], to obtain more realistic results for small transmitter–receiver distances, while approximating the conventional model at large distances. Using this model,¶¶ we concluded that $N_t^{\max}$ cannot exceed $N_t^{\boldsymbol{Q}}$, which is independent of N, but depends on A, γ, G and β. The analysis of the upper bound on $N_t^{\boldsymbol{Q}}$ in Theorem 1 has revealed that $N_t^{\boldsymbol{Q}}$ is $O(A^{\min\{\gamma/2,1\}})$ for $\gamma \neq 2$ and is $O(A/\log(A))$ for $\gamma = 2$. The analysis has also shown that $N_t^{\boldsymbol{Q}}$ is $O(\gamma^2)$ and $O(G/\beta)$.

Additionally, since the network model that we have used is quite general, our results in this paper do not only hold for the network scenarios of References [1] and [2], but also hold for networks whose nodes move with any mobility pattern or are capable of maintaining any number of simultaneous transmissions and/or receptions. Hence, we have been able to show that maximum achievable per-node end-to-end throughput is $\Theta(1/N)$, even when the mobility pattern of the nodes, the spatial–temporal transmission scheduling policy, the temporal variation of transmission powers, the source-destination pairs and the possibly multi-path routes between the nodes are optimally chosen. Furthermore, the result holds even when the communication bandwidth is divided into sub-channels of smaller bandwidth.

Moreover, our results are valid for any non-negative value of γ.‖‖ This allowed us to show that lack of attenuation and lack of space are equivalent, where $N_t^{\max}$ and $N_t^{\boldsymbol{Q}}$ cannot exceed $1+G/\beta$. Also, in these equivalent cases, λ_e and λ_m cannot exceed $W_{\max}(1+G/\beta)/(\bar{H}N)$.

We have also shown that no node can receive more than $1+G/\beta$ simultaneously successful transmissions intended for itself. This allowed us to show that $N_t^{\max}$, λ_e, and λ_m are $O(1)$ with respect to A and γ for a given N. Together with (T2.1), this also allowed us to justify that the limitation of λ_e and λ_m due to shortage of space and attenuation is more pronounced when N is large.

Finally, we have studied the implications of our results on the scalability patterns of wireless networks. We have shown that as N becomes large, unless one or more of the parameters from $W_{\max}$, γ, G/β or A grows with N, and $\bar{H}N$ is $O(W_{\max}U_\gamma)$, a desired per-node end-to-end throughput is not achievable. Regarding scalability of practical systems, we have concluded that $\bar{H}$ must be $\Theta(1)$ with respect to N. Moreover, we have concluded that A is the only feasible parameter whose growth can compensate for increasing N. Above all, we have proved that as $N \to \infty$, a desired per-node end-to-end throughput is not achievable, unless A also grows with N, and N is $O(A^{\min\{\gamma/2,1\}})$ when $\gamma \neq 2$ and is $O(A/\log(A))$ when $\gamma = 2$.

In summary, in this paper, we derived a new approach to analyze the scalability patterns of wireless networks through the use of a more general network model and we determined several necessary conditions for the scalability of this network. This was performed by considering only one of the fundamental requirements for scalability, which is the requirement of a non-vanishing per-node end-to-end throughput as the number of nodes grows large. An interesting extension of this work would be to determine the *additional* necessary conditions that result from other fundamental requirements for scalability, such as bounded end-to-end delay, bounded power consumption, bounded processing power and bounded memory consumption at the nodes.

¶¶Note that the difference between our results and the results in References [1] and [2], which concluded that $N_t^{\max}$ is $\Theta(N)$, is due to the different propagation model.

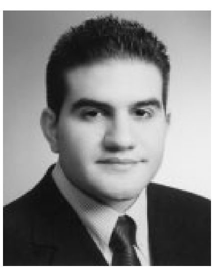
‖‖Thus, our results extend the results of References [1] and [2] that are limited to values of γ that exceed 2.

Copyright © 2004 John Wiley & Sons, Ltd.

References

1. Gupta P, Kumar PR. The capacity of wireless networks. *IEEE Transactions on Information Theory* 2000; **46**(2): 388–404.
2. Grossglauser M, Tse D. Mobility increases the capacity of wireless adhoc networks. *IEEE/ACM Transactions on Networking* 2002; **10**(4): 477–486.
3. Toumpis S, Goldsmith AJ. Capacity regions for wireless ad hoc networks. *IEEE Transactions on Wireless Communications* 2003; **2**(4): 736–748.
4. Li J, Blake C, Couto DD, Lee H, Morris R. Capacity of ad hoc Wireless Networks. In *ACM Mobicom*, Rome, Italy, July 2001.
5. Yi S, Pei Y, Kalyanaraman S. On the Capacity Improvement of ad hoc Wireless Networks using Directional Antennas. In *Proceedings of ACM MobiHoc*, Annapolis, MD, June 2003.
6. Peraki C, Servetto SD. On the Maximum Stable Througput Problem in Random Networks with Directional Antennas. In *Proceedings of ACM MobiHoc*, Annapolis, MD, June 2003.
7. Liu B, Liu Z, Towsley D. On the Capacity of hybrid wireless networks. In *IEEE INFOCOM*, San Francisco, CA, April 2003.
8. Gupta P, Kumar PR. Towards an information theory of large networks: an achievable rate region. *IEEE Transactions on Information Theory* 2003; **49**(8): 1877–1894.
9. Xie L-L, Kumar PR. A network information theory for wireless communication: scaling laws and optimal operation. To appear in *IEEE Transactions on Information Theory.*
10. Rappaport TS. *Wireless Communications, Principles and Practice*. Vol. 104. Prentice Hall: New Jersey, 1996; pp. 276–278.
11. Andersen JB, Rappaport TS, Yoshida S. Propagation measurements and models for wireless communications channels. *IEEE Communications Magazine* 1995; **33**(1): 42–49.
12. Bacceli F, Blaszczyszyn B. On a coverage range process ranging from the boolean model to the Poisson Voroni tesellation with applications to wireless communications. *Advance Application Problems* 2001; **33**(2): 293–323.
13. Dousse O, Baccelli F, Thiran P. Impact of Interferences on Connectivity in ad hoc Networks. In *IEEE INFOCOM*, San Francisco, CA, April 2003.
14. Pach J, Agarwal PK. *Combinatorial Geometry.* Wiley: New York, 1995; pp. 207–221.
15. Panik MJ. *Classical Optimization: Foundations and Extensions*. American Elsevier Publishing Co., Inc.: New York, 1976; pp. 227–306.
16. Mitrinovic DS. *Analytic Inequalities*. Springer-Verlag: New York, 1970; pp. 10–36.
17. Berkey DD, Blanchard P. *Calculus, Third Edition*. Saunders College Publishing: New York, 1992; pp 566–578.

Authors' Biographies



Onur Arpacioglu received his B.Sc. in electrical and electronic engineering degree, in 2001 with the highest honor from Bogazici University, Istanbul, Turkey. He is currently working toward the Ph.D. in the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, U.S.A. His research interests include scalability and capacity of wireless networks. His e-mail is: o.arpacioglu@cornell.edu and his URL is: http://www.people.cornell.edu/pages/oa27/.

Zygmunt J. Haas received his B.Sc. in EE, in 1979 and M.Sc. in EE, in 1985. In 1988, he earned his Ph.D. from Stanford University and subsequently joined AT&T Bell Laboratories in the Network Research Department. There he pursued research on wireless communications, mobility management, fast protocols, optical networks and optical switching. From September 1994 to July 1995, Dr. Haas worked for the AT&T Wireless Center of Excellence, where he investigated various aspects of wireless and mobile networking, concentrating on TCP/IP networks. As of August 1995, he joined the faculty of the School of Electrical and Computer Engineering at Cornell University.

Dr. Haas is author of numerous technical papers and holds fifteen patents in the fields of high-speed networking, wireless networks and optical switching. He has organized several workshops, delivered numerous tutorials at major IEEE and ACM conferences, and serves as editor of several journals and magazines, including the *IEEE Transactions on Networking, the IEEE Transactions on Wireless Communications, the IEEE Communications Magazine* and the *ACM/Kluwer Wireless Networks* journal. He has been a guest editor of IEEE JSAC issues on 'Gigabit Networks,' 'Mobile Computing Networks,' and 'Ad-Hoc Networks'. Dr Haas is a senior member of IEEE, a voting member of ACM and the chair of the IEEE technical committee on personal communications. His interests include: mobile and wireless communication and networks, personal communication service and high-speed communication and protocols. His e-mail is: haas@ece.cornell.edu and his URL is: http://wnl.ece.cornell.edu

Copyright © 2004 John Wiley & Sons, Ltd.