

Twelve Reasons not to Route over Many Short Hops

Martin Haenggi

Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN 46556, USA
E-mail: mhaenggi@nd.edu

Abstract—For multihop wireless networks, a fundamental question is whether it is advantageous to route over many short hops (short-hop routing) or over a smaller number of longer hops (long-hop routing). Short-hop routing gained a lot of support, and its proponents mainly produce two arguments: reduced energy consumption and less interference. Both arguments stem from an oversimplified analysis that is based on inaccurate channel models and neglects delay, end-to-end reliability, bias power consumption, the impact of channel coding, mobility, and routing overhead. In this paper, we shed more light on these issues by listing twelve reasons why short-hop routing is not as beneficial as it seems to be. The conclusion is that for many networks, long-hop routing is in every aspect a very competitive strategy.

I. INTRODUCTION

For certain wireless networks, such as ad hoc and sensor networks, a fundamental question is whether it is advantageous to route over many short hops (*short-hop routing* or, in the extreme case, nearest-neighbor routing) or over a smaller number of longer hops (*long-hop routing*). Recently, this debate extended to multihop extensions of WLANs [1] and multihop cellular networks [2]. Short-hop routing gained a lot of support, and its proponents mainly produce the following two arguments:

1. Energy consumption. If a long hop of distance d is divided into n hops of distance d/n , the energy benefit is often assumed to be $n^{\alpha-1}$, where α is the path loss exponent.

2. Capacity. The shorter the hops, the higher the transport capacity in an interference-limited network [3].

The first argument stems from an oversimplified analysis of the energy consumption and neglects important issues such as delay, end-to-end reliability, and bias power consumption. The second argument is only valid as long as the connectivity of the network is guaranteed; it was derived for an increasingly dense network that takes advantage of the singularity of the attenuation $d^{-\alpha}$ at $d = 0$, which may lead to the unrealistic situation that the received power exceeds the transmitter power; and it neglects delay, too. In this paper, we shed more light on these issues by listing twelve reasons why short-hop routing is not as beneficial as it seems to be. Some of the reasons have been mentioned in other work, but this is, to the

best of our knowledge, the first comprehensive collection.

Often, a *disk model*¹ is used for the analysis of wireless networks, where a transmission is either 100% successful or fails completely, depending on whether the distance is smaller or larger than the so-called transmission radius. More realistic is the *threshold model*², where a certain signal-to-noise-and-interference ratio (SINR) is needed for successful transmission. Still, for AWGN channels, the threshold model yields 0% or 100% probability and should therefore be used with great care. To get accurate results, reception probabilities should be based on bit, block, and packet errors rates, taking into account the error correction capabilities of the channel code.

We will demonstrate that by discarding the disk model and directly focusing on SINR levels, many advantages of long-hop routing become apparent.

II. NETWORK AND LINK MODEL

Part of our discussion applies to many types and classes of networks and wireless channels. However, to be concrete, we often focus on networks with random node distribution and Rayleigh fading channels.

A. Node Distribution and Generic Routing

The analytical results are derived for networks whose nodes constitute a Poisson point process in the plane. Note that for infinite networks, the Poisson point process corresponds to a uniform distribution [4], [5], and for large networks, the two distributions are equivalent for all practical purposes.

Many different routing algorithms exist for ad hoc networks [6], [7], but common to all of them is the fact that at each hop, progress shall be made towards the destination. This generic routing strategy is illustrated in Fig. 1. If the nearest neighbor within a certain sector of the source-destination axis is chosen as the next relay, this is certainly an instance of *short-hop routing*. If many nearby neighbors are skipped and a node

¹Also called protocol model [3].

²Also denoted as physical model [3].

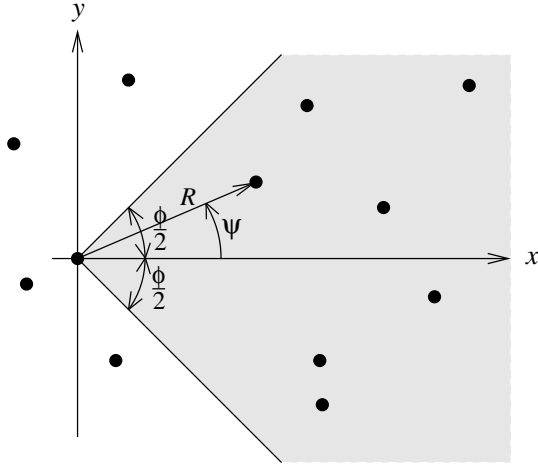


Fig. 1. Part of a Rayleigh network with the source at the origin and the x -axis pointing towards the destination node. R denotes the distance to the nearest neighbor within a sector ϕ around x , and ψ is its argument. Hence (R, ψ) are the polar coordinates of the nearest neighbor within a sector ϕ , and (X, Y) are its Cartesian coordinates.

transmits directly to a more distant neighbor, we speak of *long-hop* routing.

Note that the distance to the n -th nearest neighbor that lies within the desired sector can easily be determined from the Poisson assumption. The probability density of the distance to the n -th neighbor in a sector ϕ is [8]

$$p_{R_n}(r) = r^{2n-1} \left(\frac{\phi}{2}\right)^n \frac{2}{(n-1)!} e^{-r^2\phi/2} \quad (1)$$

Since p_{R_n} is a Rayleigh distribution for $n = 1$, it can be considered a generalized Rayleigh distribution. Similarly, in one dimension, the Erlang distribution is a generalized exponential distribution. So, the transition from one dimension to two dimensions simply entails a multiplication by $r\phi$ (that comes from the inner derivative of the exponential part) in the distributions of the node distances.

B. Link Model

Although some of the reasons listed in the next section also pertain to the disk model, we will mostly use a Rayleigh block fading channel model. Assuming that a certain SINR level Θ is required for successful reception (threshold model), the reception probability of successful packet reception can be expressed as [8]

$$p_r = \underbrace{\exp\left(-\frac{\Theta N}{P_0 d_0^{-\alpha}}\right)}_{p_r^N} \cdot \underbrace{\prod_{i=1}^k \frac{1}{1 + \Theta \frac{P_i}{P_0} \left(\frac{d_0}{d_i}\right)^\alpha}}_{p_r^I}, \quad (2)$$

where N is the noise power, P_0 the transmit power of the desired transmitter, P_i , $i > 0$ the transmit powers of the k interferers (who transmit in the same time slot), d_0 the distance from the receiver to the desired transmitter, and d_i , $i > 0$, the

distance to the interfering transmitters. α is the large-scale path loss exponent. In the expression (2), we can identify two parts: p_r^N , which is the reception probability in an interference-free network since it only depends on the noise, and p_r^I , which is the reception probability in a noise-free network since it only depends on the interference. This property of the Rayleigh fading channel enables the separate analysis of noise and interference effects.

III. THE TWELVE REASONS

Clearly, the most compelling reason against short-hop routing is the end-to-end delay. However, we do not consider delay in itself an argument for long-hop routing, since energy and delay can be traded off against each other. So, for a fair comparison, both schemes shall satisfy the same energy and delay requirements. We may fix the energy consumption and search for the protocol with the smallest delay, or we may impose a delay constraint and determine which protocol consumes the least amount of energy.

A. Interference

According to [9], “It is unclear whether more interference is caused by a single transmission at higher power or multiple transmissions at lower power”. Indeed, a shorter transmission at higher power may permit more efficient reuse of the communication channel. If the total radiated energy (product of power and duration) of is a good indicator for interference, this boils down to an energy consumption problem. However, it must not be forgotten that *the SIR does not depend on absolute power levels*. If all nodes scale their power by the same factor $q > 1$, all the SIR levels remain constant, but the SINR levels will increase, which can be easily seen from (2). So, increasing all transmit power levels does not have a negative impact on any packet reception probability in the network (on the contrary), in stark contrast to what is predicted by the disk model. This indicates that *long-hop transmission does not inherently cause more interference*.

B. End-to-end Reliability

Under the disk model, reception probabilities are either 100% or 0%. If every receiver is in its desired transmitter’s disk, the end-to-end reliability is always 100%, which is clearly not realistic, since packet errors or bit errors accumulate. In the Rayleigh block fading case, the end-to-end reception probability p_n over an n -hop route follows from (2):

$$p_n = \prod_{i=1}^n e^{-\Theta/\bar{\gamma}_i} = e^{-\Theta \sum_{i=1}^n \frac{1}{\bar{\gamma}_i}} \quad (3)$$

where $\bar{\gamma}_i$ denotes the mean SNR at link i . So, to achieve a desired end-to-end reliability with short-hop routing, the relay nodes need to transmit at a higher power. This compensates, at least partially, for the loss in SNR.

Consider the case of a one-dimensional chain of equidistant nodes with distance d (this is the optimum scenario for short-hop routing). Let E_0 be the energy required for a transmission over distance d with probability P_{EE} , i.e., $E_0 := -d^\alpha \Theta N / \ln P_{EE}$. Covering the total distance in one single hop requires an energy of $E_1 = n^\alpha E_0$. In the multihop case with n hops, a reception probability $p_r = \sqrt[n]{P_{EE}}$ is required at each hop. Since $\ln P_{EE} = n \ln p_r$, the total energy in this case is $E_n = n \cdot n E_0$. So, for $\alpha = 2$, there is no benefit in multihop routing.

If the same end-to-end delay is permitted, a number of transmission (at lower power) are possible in the long-hop case. For the block fading channel, this results in a time diversity benefit, and the energy advantage of short-hop routing vanishes also for higher path loss exponents, in particular for high end-to-end reliabilities or when channel state information is available at the transmitter.

C. Capacity and Channel Coding

For optimum coding in AWGN channels, a change from a nominal capacity $C_0 := \log_2(1 + P_0/N)/2$ to C_0/q results in an energy consumption of $E(q) = qP(q) = qN(2^{2C_0/q} - 1)$, which is strictly monotonically decreasing in q .³ Note that $C_0/q =: R$ is the information theoretic rate (bits/symbol), and that the gain from using longer codes is higher for higher rates.

Assume that a (normalized) distance 1 is covered by an n -hop route for some small n , say $1 \leq n \leq 5$. In this case, spatial reuse is not possible, so a simple TDMA MAC scheme will perform optimally. Since there is no interference in this case, the analysis is based on AWGN channels. Let R denote the bandwidth-normalized end-to-end rate for the n -hop route. We find

$$R = \frac{1}{n} \log_2 \left(1 + \frac{E_s}{N_0} n^\alpha \right), \quad (4)$$

where $E_s = E_b R$ is the energy per symbol and N_0 is the noise spectral density. By dividing the distance into hops of length $1/n$, the SNR increases by n^α . On the other hand, the end-to-end rate is reduced by a factor of n . Conversely, the n -hop scheme needs to transmit at a per-hop rate of nR to achieve the same end-to-end rate. The question is which is the optimum n given a certain desired end-to-end rate. This problem is addressed in some detail in [10]. Here we note that since the rate loss in (4) is linear while the gain from the increased SNR is only proportional to $\log_2(n)$, there exists a certain rate R_n at which the bit-energy-to-noise ratio E_b/N_0 is the same for n -hop and $n+1$ -hop routing. Above R_n , n -hop routing performs better, and below R_n , $(n+1)$ -hop routing is better. Focusing on the case of single-hop and two-hop routing, we find from (4) that $R_1 = \log_2(2^\alpha - 1) < \alpha$. So, as a simple

³The increase in transmission length is only linear in q , while the power can be reduced exponentially.

rule of thumb, we can say that whenever the desired rate is higher than the path loss exponent, single-hop routing achieves the highest capacity.

More generally, it can be shown that for $R \rightarrow 0$, there exists an asymptotic per-hop spectral efficiency

$$S := \lim_{R \rightarrow 0} R n_{\text{opt}}(R) \quad (5)$$

that only depends on α . It is given by

$$S = \frac{W(-\alpha e^{-\alpha}) + \alpha}{\ln 2} \lesssim \frac{\alpha(1 - e^{-\alpha})}{\ln 2}, \quad (6)$$

where $W(\cdot)$ denotes the (principal branch of the) Lambert W function and the bound stems from the first order Taylor expansion. The primary use of this asymptotic spectral efficiency is to determine the optimum hop number: It can be shown that with good accuracy

$$n_{\text{opt}}(\alpha, R) \approx [S(\alpha)/R], \quad (7)$$

where $[x]$ denotes the nearest positive integer to x . Using the upper bound (6), we find

$$n_{\text{opt}}(\alpha, R) \leq \left\lceil \frac{\alpha(1 - e^{-\alpha})}{R \ln 2} \right\rceil, \quad (8)$$

which conveniently yields the optimum number of hops for any R and α .

If channel coding is taken into account, multihop is further penalized due to the necessary encoding and decoding at each hop.

D. Total Energy Consumption

It is often assumed that a reduction of the *transmit energy* yields a proportional reduction of the total energy consumption. Even without taking into account *receive energy*, this is not true for any practical power amplifier. In particular in low-power transceivers, the local oscillators and bias circuitry will dominate, so that short-hop routing does not yield any energy benefit if a more distant relay node can be reached with sufficient reliability [11]. For random networks, relatively high peak power levels are necessary to keep the network connected [12], and short-hop routing would require a substantial backoff on the average, resulting in poor power efficiency.

E. Path Efficiency in Random Networks

Routes in random networks cannot follow straight lines. The path efficiency, defined as the ratio of Euclidean distance of the end nodes and the travelled distance, is higher if longer hops are used. Consider the generic routing strategy in Fig. 1. For nearest-neighbor routing in a sector ϕ the expected path efficiency η for a long connection is $\mathbb{E}[\cos \Psi]$, so

$$\eta(\phi) = \mathbb{E}[\cos \Psi] = \frac{2}{\phi} \int_0^{\phi/2} \cos \psi \, d\psi \quad (9)$$

$$= \frac{2}{\phi} \sin \left(\frac{\phi}{2} \right) \approx 1 - \frac{\phi^2}{24}, \quad (10)$$

where the approximation is the second-order Taylor expansion. Now assume that this nearest-neighbor path has been established, and that, instead of routing through every node, a long-hop strategy is applied where only every n -th node is used as a relay. With increasing n , the argument Ψ from the origin to the n -th node tends to be Gaussian distributed with variance $V(\phi, n)$. Since the support of the pdf of Ψ is always $[-\phi/2, \phi/2]$,⁴ the variance decreases inversely proportional to n , i.e., $V(\phi, n) \approx \frac{\phi^2}{12n}$. So, for large n , we get⁵

$$\eta_n(\phi) = \mathbb{E}[\cos \Psi] \quad (11)$$

$$\approx \frac{\sqrt{6n}}{\phi\sqrt{\pi}} \cdot \int_{-\infty}^{\infty} \cos \psi e^{-6n(\frac{\psi}{\phi})^2} d\psi \quad (12)$$

$$= e^{-\frac{\phi^2}{24n}} \quad (13)$$

$$\approx 1 - \frac{\phi^2}{24n}, \quad (14)$$

which shows that the path efficiency grows monotonically with n and goes to 1 for large n . The energy penalty caused by deviations from the optimum path is

$$\frac{1}{\eta_n(\phi)^\alpha} \approx 1 + \frac{\alpha\phi^2}{24n}. \quad (15)$$

This shows that the energy penalty can be made independent of α , if only every $n = \alpha$ -th node is used.

F. Sleep Modi or Cooperation

If neighboring nodes are not used as relays, they can either be put to sleep, or they can assist the transmission by cooperation [13], [14] or retransmission (e.g., if an ACK packet is not received by the source).

G. Routing Overhead and Route Maintenance

In [9], it is pointed out that (when we replace a larger number of short hops by a smaller number of long hops) “It is far from clear what happens to the overall transmission energy, since to implement a nearest-neighbor policy, significantly augmented overhead control traffic will be required to coordinate the establishment of the routing paths and access control protocols across the entire network.”

In a first order approximation, the control traffic for routing and route maintenance is proportional to the number of nodes in the route. Also, the probability of a route break due to energy depletion and node failure clearly increases with the number of nodes involved, as well as the memory requirements for the routing tables.

⁴Hence, after every convolution, the support needs to be scaled, which results in a reduction of the variance.

⁵The Gaussian approximation is very accurate even for small n . For $\phi = \pi/2$ and $n = 1$, e.g., the precise value is $2\sqrt{2}/\pi \approx 0.9003$, whereas this approximation yields $e^{-\pi^2/96} \approx 0.9023$, so the error is only 0.2%. The second order Taylor expansions are identical, even for $n = 1$.

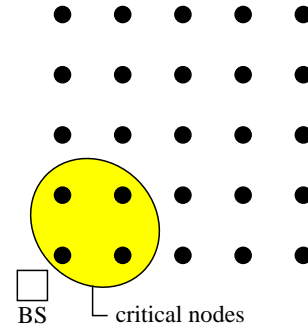


Fig. 2. The *critical area* of a sensor network is determined by the nodes that can reach the base station in a single hop. If it consists of only a few nearest neighbors of the base station, these nodes will severely limit the lifetime of the network since they have to relay the whole traffic.

H. Route Longevity in Mobile Environments

The SNR of short-hop routes is more quickly affected by moving nodes. For example, if a node at distance 1 moves away by 1 unit, the SNR change is 2^α , which causes the link to break (unless an unreasonably high SNR margin is applied). On the other hand, if a relay node is 3 units away moves by the same distance, the SNR change is only $(4/3)^\alpha$, which can probably be tolerated.

In general, the lifetime of a link is proportional to its (original) length. To show this, assume that routes are established with a certain SNR margin M_Θ , such that for a link of initial distance d , we have $\frac{P_0}{d^{\alpha N}} = \Theta M_\Theta$. At the critical distance $d + \Delta d$ when the link breaks, we have $\frac{P_0}{(d + \Delta d)^{\alpha N}} = \Theta$. Solving for Δd , we find

$$\Delta d = d(M_\Theta^{1/\alpha} - 1). \quad (16)$$

For a given mobility pattern, Δd determines the lifetime of a link, so indeed longer links live proportionally longer.

I. Traffic Accumulation and Energy Balancing

For certain multihop networks such as sensor networks or multihop cellular networks, traffic accumulation around a base station (BS) or access point is a big problem. With strict short-hop routing, the relaying burden cannot be distributed among a high enough number of nodes, leading to a *critical area* around the BS whose nodes suffer from a short lifetime (see Fig. 2) The more nodes can reach the BS directly, the better distributed the load can be [15].

J. Variance in Hop Length

In random networks, due to the variance in hop length, the variance in energy consumption is large when nearest-neighbor schemes are used, causing substantial imbalance in the energy consumption. Assuming that a power control scheme is used that adapts the transmit power to maintain a certain reception probability p_r^N (see (2)). For an n -hop route, we are interested in the power consumption at the node

that has to transmit over the largest distance, since this node determines the lifetime of the route. From (1) we know that the distance to the nearest neighbor is Rayleigh distributed. So, for a path loss exponent $\alpha = 2$, the power level is given by an exponential distribution, and we need to determine the expected maximum of n exponentially distributed random variables. The same problem appears in selection combining of multiple transmission in fading channels, so we can apply the result from [16, p.316] that this maximum is given by the harmonic sum, *i.e.*,

$$\mathbb{E}[\max\{R_1, \dots, R_M\}^2] = \sum_{k=1}^M \frac{1}{k} \approx \ln(M) + \gamma_{\text{em}}, \quad (17)$$

where $\gamma_{\text{em}} \approx 0.577$ is the Euler-Mascheroni constant. The bound $\ln(M) + \gamma_{\text{em}}$ is tight already for moderate M , so the expected maximum $\mathbb{E}[\max\{R_1, \dots, R_M\}^2]$ is sharply lowerbounded by $\mathbb{E}[R^2](\ln M + \gamma_{\text{em}})$. For $\alpha \geq 2$, with Jensen's inequality, we have $\mathbb{E}[(X^2)^{\alpha/2}] \geq \mathbb{E}[X^2]^{\alpha/2}$ and thus

$$\begin{aligned} \mathbb{E}[\max\{R_1, \dots, R_M\}^\alpha] &\geq (\mathbb{E}[\max\{R_1, \dots, R_M\}^2])^{\alpha/2} \\ &\gtrsim (\mathbb{E}[R^2](\ln M + \gamma_{\text{em}}))^{\alpha/2}, \end{aligned} \quad (18)$$

or, since $\mathbb{E}[R^2] = (4/\pi)\mathbb{E}[R]^2$,

$$\mathbb{E}[\max\{R_1, \dots, R_M\}^\alpha] \gtrsim \left(\frac{2}{\sqrt{\pi}}\mathbb{E}[R]\right)^\alpha (\ln M + \gamma_{\text{em}})^{\alpha/2}. \quad (19)$$

For $\alpha = 1$ (the distance itself), we have the *upper* bound

$$\mathbb{E}[\max\{R_1, \dots, R_M\}] \lesssim \sqrt{\mathbb{E}[R^2](\ln M + \gamma_{\text{em}})}. \quad (20)$$

This shows that the maximum energy consumption over M nearest-neighbor hops grows with at least $\ln M$ in networks with random node distribution. If longer hops are permitted, the distances and thus the power consumption can be better equalized among the nodes in a route.

K. Bounded Attenuation

As pointed out in the introduction, a path loss model with a singularity at distance $d = 0$ is not realistic for networks with high density. Clearly, there is a bound on the received power. If we assume that this bound is achieved for distances $0 < d < R$, then there is no benefit of using shorter hops than R , since the interference remains constant and the received power does not increase if the distance is decreased further. This problem is particularly relevant in dense networks [17].

L. Multicast Advantage

So far, we have only addressed unicast routing. For multicast, other tradeoffs between short-hop and long-hop routing exist. In particular, as discussed in [9], it is often advantageous for a source to transmit at high power levels to reach a maximum number of nodes in the multicast group.

IV. CONCLUSION

We have listed twelve compelling reasons why the tradeoff between routing over many short hops and routing over fewer longer hops is not as clear as it is often assumed. Not all reasons apply to all types of networks, of course, but several of them will be relevant for most networks. The conclusion is that *routing as far as possible is a very competitive strategy* in many cases. Conversely, from a design perspective, the peak transmit power should be chosen such that a node can reach well beyond nearest neighbors.

REFERENCES

- [1] J. Jun and M. L. Sichitiu, "The Nominal Capacity of Mesh Networks," *IEEE Wireless Communications*, vol. 5, pp. 8–14, Oct. 2003.
- [2] Y.-D. Lin and Y.-C. Hsu, "Multihop Cellular: A New Architecture for Wireless Communications," in *IEEE INFOCOM'00*, vol. 3, pp. 1273–1282, 2000.
- [3] P. Gupta and P. R. Kumar, "The Capacity of Wireless Networks," *IEEE Transactions on Information Theory*, vol. 46, pp. 388–404, Mar. 2000.
- [4] E. S. Sousa and J. A. Silvester, "Optimum Transmission Ranges in a Direct-Sequence Spread-Spectrum Multihop Packet Radio Network," *IEEE Journal on Selected Areas in Communications*, vol. 8, pp. 762–771, June 1990.
- [5] R. Meester and R. Roy, *Continuum Percolation*. Cambridge University Press, 1996.
- [6] E. M. Royer and C.-K. Toh, "A Review of Current Routing Protocols for Ad-Hoc Mobile Wireless Networks," *IEEE Personal Communications*, vol. 6, pp. 46–55, Apr. 1999.
- [7] C. E. Perkins, ed., *Ad Hoc Networking*. Addison Wesley, 2000. ISBN 0-201-30976-9.
- [8] M. Haenggi, "On Routing in Random Rayleigh Fading Networks," *IEEE Transactions on Wireless Communications*, 2004. Accepted for publication. Available at <http://www.nd.edu/~mhaenggi/routing.pdf>.
- [9] A. Ephremides, "Energy Concerns in Wireless Networks," *IEEE Wireless Communications*, vol. 9, pp. 48–59, Aug. 2002.
- [10] M. Sikora, J. N. Laneman, M. Haenggi, D. J. Costello, and T. Fuja, "On the Optimum Number of Hops in Linear Ad Hoc Networks," in *IEEE Information Theory Workshop (ITW'04)*, (San Antonio, TX), Oct. 2004. Submitted for publication.
- [11] M. Haenggi, "The Impact of Power Amplifier Characteristics on Routing in Random Wireless Networks," in *IEEE Global Communications Conference (GLOBECOM'03)*, (San Francisco, CA), Dec. 2003. Available at <http://www.nd.edu/~mhaenggi/globecom03.pdf>.
- [12] F. Xue and P. R. Kumar, "The Number of Neighbors Needed for Connectivity of Wireless Networks," *Wireless Networks*, 2004. To appear. Available at http://black.csl.uiuc.edu/~prkumar/ps_files/connect.pdf.
- [13] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Transactions on Information Theory*. Accepted for publication. Available at: <http://www.nd.edu/~jnl/pubs/it2002.pdf>.
- [14] M. Haenggi, "Analysis and Design of Diversity Schemes for Ad Hoc Wireless Networks," *IEEE Journal on Selected Areas in Communications*. Accepted for publication. Available at http://www.nd.edu/~mhaenggi/jsac_adhoc.pdf.
- [15] M. Haenggi, "Energy-Balancing Strategies for Wireless Sensor Networks," in *IEEE International Symposium on Circuits and Systems (ISCAS'03)*, (Bangkok, Thailand), May 2003. Available at <http://www.nd.edu/~mhaenggi/iscas03.pdf>.
- [16] W. C. Jakes (Ed.), *Microwave Mobile Communications*. IEEE Press, 1993. ISBN 0-7803-1069-1.
- [17] O. Dousse and P. Thiran, "Connectivity vs Capacity in Dense Ad Hoc Networks," in *IEEE INFOCOM*, (Hongkong), Mar. 2004.