Analysis of Wireless Ad-Hoc and Sensor Networks in Finite Regime

Hossein Pishro-Nik and Faramarz Fekri

***Abstract*— In the past, many analytic results for wireless networks have been reported for the case where the number of nodes n in the network tends to infinity (large-scale networks). These include connectivity, coverage, and capacity. These results have not been extended for small or moderate values of n, although in many practical networks n is not very large. In this paper, we first show that previous asymptotic results provide poor approximations for the finite networks (small-scale networks). We then aim to develop a framework to analytically study network properties without assuming that n is large. We provide a set of differences between small-scale and large-scale analysis. We consider wireless networks in which the location of the nodes is random. We study routing algorithms, coverage, connectivity and capacity of finite wireless networks. We provide easily computable expressions for different network properties. With validation from simulations, we show that these analytic expressions give very good estimates of these quantities for finite wireless networks. Our investigation suggests that the small-scale networks posses unique characteristics that require a new framework for analysis and design.**

***Index Terms*— Wireless Networks, Small-Scale Networks, Graph theory Connectivity, Coverage, Capacity.**

I. Introduction

In the past, many analytic results on the connectivity, coverage, and capacity of wireless ad-hoc and sensor networks have been obtained. In almost all of the results, it is assumed that the number of nodes n in the network tends to infinity (large-scale networks). In other words, these results are asymptotic. Asymptotic results are very important for two reasons. First, they give us good estimates for large-scale networks. Second, they show some fundamental trade-offs in the network. However, in many practical wireless networks the number of nodes may be limited to a few hundred (small-scale/finite networks). As it is shown in this paper, the asymptotic results cease to be valid for these networks. Thus, it is very crucial from practical point of view to analyze finite networks. These analytic results will essentially help us to understand, design, and analyze practical wireless networks, and also to design more suitable communication protocols.

To clarify, for example, consider capacity analysis of wireless networks which has been studied extensively (e.g., in [1]–[7]). Today we have good understanding of scaling laws in capacity of wireless networks. However, suppose we need to design a wireless sensor network consisting of a hundred sensor nodes. Some fundamental question are as follows. What is the transport capacity? What are the information theoretic and the MAC layer capacities? How do network parameters such as number of nodes and their communication radius affect these capacities? Unfortunately, the available asymptotic results fail to give answers to these questions. Similar questions are remained unanswered for other properties of the network such as connectivity, coverage, etc.

The question that arises here is, can we do small-scale analysis? We recognize some obstacles as follows. First, in large-scale networks we can use asymptotic estimates that make the analysis much simpler. These estimates are not available in small-scale analysis. Thus, small-scale analysis is usually more difficult. Second, even if we can perform the small-scale analysis, we usually obtain very complicated formulas that are not very useful practically. In this paper, we want to circumvent these problems and provide guidelines for small scale-analysis. We assume the reader is, to some extent, familiar with large-scale (asymptotic) analyses. The main goal of the paper is to initiate the small-scale analysis of wireless sensor and ad hoc networks. Such analyses can be very useful in analyzing and evaluating communication and security protocols for practical sensor and ad hoc networks and is completely overlooked in the literature. To the best of our knowledge, this is the first work to analytically and systematically study this issue.

The main idea is the following. First, for clarity, by small-scale (finite) networks we mean networks of size between $n = 20$ and $n = 2000$, which includes many practical wireless sensor and ad hoc networks. The first key point is to aim at simple and very good approximations instead of trying to find complicated exact formulas. To do so, we first consider the asymptotic analysis. In any asymptotic analysis, a set of asymptotic estimates are used. Some of these estimates are still good for small-scale networks, while others are not. We identify those who are not valid and replace them with better estimates. Specifically, in this paper we list a few important differences between small-scale and large-scale analysis. Some of these differences, such as the field-shape effect, are specific to random geometric graphs while others apply to all finite and asymptotic systems. Thus, the general method is that we look at any asymptotic analysis and identify the estimations that are not valid for finite networks and replace them with more accurate estimates. However, this must be done carefully, in order to obtain simple and easily computable formulas at the end. As it is mentioned above, exact expressions for network quantities are usually very complicated. Thus, we attempt to provide easily computable estimates for those quantities.

H. Pishro-Nik is with the Electrical and Computer Engineering Department of the University of Massachusetts, Amherst, MA, USA (email:pishro@ecs.umass.edu).

F. Fekri is with the School of Electrical and Computer Engineering of the Georgia Institute of Technology, Atlanta, Georgia, USA (email:fekri@ece.gatech.edu).

This work is supported by National Science Foundation under grants CCF-0728970 and CCF-0728772.

978-1-4244-1777-3/08/$25.00 © IEEE

In this paper, we consider fundamental network properties that affect routing algorithms and reliability of the wireless networks. Specifically, we study coverage, connectivity, capacity and analysis of routing algorithms. We give several results pertaining to these properties. For example, as an important property we want the network to be connected. More generally we may need k-connectivity. For multi-path routing using k disjoint paths between different nodes, we need the network to be k-connected. Moreover, k-connectivity is related to reliability of networks against node and link failures and adversaries. K-connectivity is also desirable in networks with sleeping sensors. In the past, many authors have studied connectivity and k-connectivity for large-scale networks. These results are asymptotic and obtained assuming that the number of nodes tends to infinity. Here we show that these results are not very useful for finite networks. We provide a very simple formula for k-connectivity probability of wireless network and show that the formula is very precise. Due to the space limitation, lengthy proofs and discussions are omitted here, and have been included in the longer version of the paper available at on the first author's website.

Here, we consider a wireless network that consists of n nodes and assume that the nodes are placed on a plane based on a given probability distribution. For example, in wireless sensor networks it is usually assumed that the nodes are randomly and uniformly deployed over a given field [8]. We assume that each node has a finite and fixed communication radius. Two nodes are connected (i.e., can communicate with each other) if they are within communication range of each other. Throughout the paper, we assume $\mathcal{B}(\mathbb{R}^2)$ is the Borel $\sigma-$algebra on $\mathbb{R}^2$ and m is the Lebesgue measure on $\mathcal{B}(\mathbb{R}^2)$. Note that we just mention measure theoretic definitions to take care of technicalities, and it is not necessary for the reader to be familiar with them. The reader can simply assume that for a set F in $\mathbb{R}^2$, $m(F)$ is the area of F. $B(\overline{X}, R)$ is the closed ball with radius R centered at $\overline{X}$ in $\mathbb{R}^2$. $S(\overline{X}, L)$ is the closed square with side L centered at $\overline{X}$ in $\mathbb{R}^2$. In particular $S_0 = S(\overline{O}, 1)$ is the closed square with unit area centered at the origin. If u and v are two nodes of a network located in $\mathbb{R}^2$, then $d(u, v)$ is the Euclidean distance between the location of the points. For any set $F \in \mathcal{B}(\mathbb{R}^2)$ we define $\nu(F) = m(F \cap S_0)$. Clearly, ν defines a measure on $\mathcal{B}(\mathbb{R}^2)$. Let $\mathcal{E}_n$ be an event depending on a parameter n. We say that $\mathcal{E}_n$ holds asymptotically almost surely if Pr$\{\mathcal{E}_n\}$ tends to 1 as $n \to \infty$.

The remainder of the paper is structured into several parts. The next section provides an overview of the work related to our study. Section III establishes the formulation and preliminaries of the problem we have considered. In Section IV, we justify the need for small-scale analysis developed in this paper. In Section V, we investigate the fundamental properties of small-scale analysis. We study coverage, connectivity, capacity, and routing algorithms of finite wireless networks. Finally, Section VI concludes the paper.

II. Related Work

Related problems have been studied in the context of random graph theory [9], continuum percolation and geometric probability [10], [11], and the study of wireless network graphs [2], [12]–[20]. In random graph theory, the model $G(n, p)$ is extensively studied, in which edges appear in a graph of n vertices with probability p independent of each other. In continuum percolation theory, usually infinite graphs on $\mathbb{R}^d$ are studied. Finally, in geometric probability and the study of graphs of wireless networks, large-scale graphs over the plane are usually studied.

In [12], the connectivity of large-scale wireless networks is studied. In [18], [19], and [21], k-connectivity of wireless networks has been investigated. In [18], k-connectivity was explored in the context of fault-tolerant networks. In [19] authors studied the asymptotic critical transmission radius for k-connectivity and asymptotic critical neighbor number for k-connectivity in wireless networks. In [21], we extended connectivity and k-connectivity for large-scale sensor networks. In that paper, we specifically considered the effects of node and link failures and the distribution function of the nodes on connectivity properties of sensor networks. The connectivity in ad hoc and hybrid networks was studied in [22]. In [23], trade-off between connectivity and capacity of dense networks was examined. In particular, the effect of the attenuation function on network properties was considered. Medium access (MAC) layer capacity of wireless ad hoc networks has been studied in [1]. The transport and information theoretic capacity has been investigated extensively, for example see [2]–[7]. In [20], the general concept of k-coverage was studied. The grid model for sensor networks has also been investigated. In particular, connectivity, coverage, and diameter of sensor grids were studied in [16]. In [20], the k-coverage problem for sensor grids and other deployment methods was considered and necessary and sufficient conditions were obtained. In [24], a different model for grids was examined and its asymptotic connectivity and lifetime were explored. However, almost all previous analytical results are asymptotic since they consider large-scale networks.

Analysis of wireless networks with modest number of nodes, which does not involve asymptotic results, has generated a lot of interest in the recent past [25]–[29]. In [25], the authors investigated the problem of connectivity in ad hoc networks of finite size. The probability of connectivity was analyzed for one-dimensional networks (line networks), and the result was extended to find bounds for the networks in two dimensions. The authors considered a network consisting of nodes uniformly distributed in the deployment area (a line network), where the locations of the nodes are independent of each other. Using probabilistic methods, they obtained the exact formulation for the probability of connectivity. As expected, they observed that in finite-scale networks, the probability of connectivity does not exhibit threshold effects, unlike in the asymptotic case. The author of [26] presented corrections and extensions to [25]. It is noted that both of the above cases considered a line network, and the extension to two-dimensional networks was achieved by obtaining a loose bound using the results from the former case. In [29], the authors also consider the line network and obtain connectivity results for the one-dimensional networks. It should be noted that, as it will be clear from the next discussions, the main challenges in finite analysis arise in two dimensional case. In [27], mobility and more realistic models were examined. The authors obtained an analysis of connectivity for both finite and asymptotic cases. They obtained a mobility model based on a stationary AR-1 process. One step conditional connectivity probability was found first, and an extension to the k-step case was made. Further, the author also obtained the probability mass function of the first passage disconnectivity time given that the network is initially connected. In [28], some simple local network characteristics such as link probability and average node degrees are studied. the paper also obtains formulas for average covered area. In this paper, we are interested in global properties such as connectivity, coverage probability, capacity, throughput, and lifetime. These properties are much more difficult to analyze than the properties studied in [28], and are also more relevant to the practical performance of wireless networks.

Though all the above works provide the necessary insight into analysis of finite-scale networks, the results presented, unlike the compact closed-form results in large-scale scenarios such as in [11] or [30], involve complicated series and products. This implies that one cannot get a fair insight into the dependency of network properties on design parameters (scaling laws in finite domain). Further, there exist a number of works in literature which are independent of the size of the network, but use algorithmic and simulation approach to solve such problems.

III. Preliminaries

Wireless networks are sometimes modeled with the probability space of graphs that we represent with $g(n,r) = g(n,r(n))$. In this model, it is assumed that n nodes are uniformly and randomly distributed over $S_0 = \overline{S(\overline{O},1)}$. If two nodes u and v satisfy $d(u,v) \leq r(n)$, then the edge $\{u,v\}$ belongs to edges of the graph. A more general model is the model $g(n,r,p)$, in which two nodes are connected with probability $0 < p \leqslant 1$ if their distance is less than r. In this model p models link failures that are common in wireless networks. Asymptotic properties of $g(n,r)$ have been studied extensively. Here we are interested in these properties when n is not necessarily large. It is worth noting that the assumption that the nodes are distributed on S_0 is made for simplicity. These arguments can easily be generalized to other models for the deployment region.

Another generalization is given by $g(n,r(n),f_{XY})$, which is defined as follows. Let X and Y be absolutely continuous random variables with continuous joint density function $f_{XY}(x,y)$ satisfying $f_{XY}(x,y) > 0$ for all $(x,y) \in S_0 = \overline{S(\overline{O},1)}$, and $f_{XY}(x,y) = 0$ otherwise. A graph in $g(n,r,f)$ has n nodes and is generated as follows. For any node v, its location (X,Y) is chosen according to $f_{XY}(x,y)$ independently from other nodes. If two nodes u and v satisfy $d(u,v) \leq r(n)$, then the edge $\{u,v\}$ belongs to edges of the graph. Here for simplicity, we restrict ourselves to the model $g(n,r(n))$ and $g(n,r,p)$ (i.e., when $f_{XY}(x,y) = 1_{\{(x,y)\in S_0\}}$, the uniform distribution of nodes). Again, all the arguments can be easily extended to a general density function $f_{XY}(x,y)$. For the purpose of analysis, we divide the square S_0 to different parts shown in Fig.1.

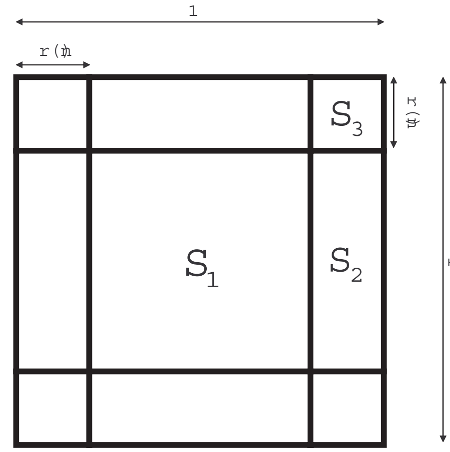

Fig. 1. The field S_0 and its subdivisions S_1,S_2, and S_3.

Finally, we consider the following definition for Poisson processes. For a point process χ on $\mathbb{R}^2$ and a Borel set A, let $\chi(A)$ be the number of points of the process in A. The point process χ_λ is said to be a Poisson process with density $\lambda > 0$ if [11]

- For mutually disjoint Borel sets $A_1, A_2, ..., A_k$, the random variables $\chi(A_1), ..., \chi(A_k)$ are mutually independent.
- For any bounded Borel set $A \in \mathcal{B}(\mathbb{R}^2)$ and for every $k \geq 0$, we have

$$\Pr\{\chi(A) = k\} = e^{-\lambda m(A)} \frac{\lambda^k (m(A))^k}{k!}. \quad (1)$$

IV. Motivation for Small-Scale Analysis

In this section, we present some evidence to show that previous asymptotic results diverge significantly from actual values for finite networks. To show this, we consider connectivity. We first provide the asymptotic probability of disconnectivity for $g(n,r,p)$ and compare it to simulation results. Using this, we conclude that the asymptotic results fail to provide an acceptable estimate of real probability of disconnectivity for small-scale networks. The following result is proved in [12], where a slightly different model is considered. However, the results can be trivially extended to $g(n,r)$.

Theorem 1: (Gupta and Kumar 1998) Let $c_n = n\pi r^2 - \log(n)$, then $g(n,r)$ is connected with high probability if $\lim_{n\to\infty} c_n = \infty$. On the other hand, if $\lim_{n\to\infty} c_n = c < \infty$ then for large n, $g(n,r)$ is disconnected with a strictly positive probability $1 - p_{asymp}(c)$.

This theorem states that if $\lim_{n\to\infty} c_n = c < \infty$, the network connectivity probability will be bounded away from one. In fact, $p_{asymp}(c)$ is the limit for the probability that the network is connected when n goes to infinity. Although, there is vast literature on the asymptotic analysis of connectivity properties of wireless networks, we were unable to find a reference that actually gives a formula for $p_{asymp}(c)$. Thus, here we compute $p_{asymp}(c)$.

Theorem 2: Let $c_n = np\pi r^2 - \log(n)$, and $\lim_{n\to\infty} c_n = c < \infty$, then the probability that $g(n, r, p)$ is connected , $p_{asymp}(c)$, satisfies

$$p_{asymp}(c) = \lim_{n\to\infty} p_{asymp}(c, n) = e^{-e^{-c}}. \quad (2)$$

Proof: For proof, please see the longer version of the paper available at http://www.ecs.umass.edu/ece/pishro/. ■

Therefore, asymptotically, the probability that $g(n, r, p)$ is connected is given by $p_{asymp} = e^{-ne^{-np\pi r^2}}$. We now show that the above asymptotic connectivity formula results in a very bad estimate of disconnectivity probability for small-scale networks. However, in the next sections, we will confirm that our small-scale analysis gives a very good estimate for this quantity.

In Figure 2, we compare the probability of having a disconnected graph for $n = 100$ and $p = 1$ derived by exhaustive simulations and the asymptotic result. In the figure, the probability of disconnectivity is shown as a function of r, the communication radius. The experiment shows that these results may differ by 10 orders of magnitude. This illustrates that the asymptotic method fails to provide a good approximation for small- scale networks.

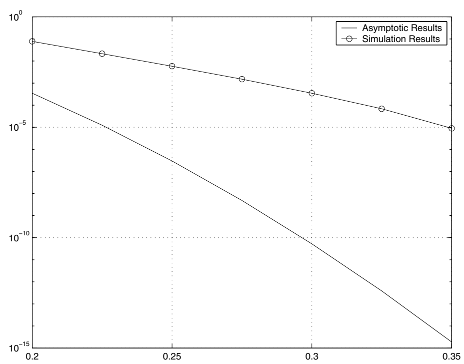

Fig. 2. Comparison of asymptotic results with the small scale simulation results for the probability of disconnectivity of $g(n = 100, r)$.

So far we showed that the asymptotic analysis may fail badly for finite networks. Now if we ought to use finite-scale analysis, what kind of formulation would be helpful? To answer, let us now elaborate on the important requirement we mentioned earlier. Namely, in small-scale analysis we need to find simple and easily computable formulas. The rationale behind this is as follows. First, in analytic results we usually need formulas that help us to understand the effects of different parameters. A complicated formula usually reveals little about those effects. Second, sometimes, exact formulas are computationally infeasible. To show this, let us again consider connectivity.

For n points $\overline{X}_1, \overline{X}_2, ..., \overline{X}_n$ on the plane, let the graph $g(\overline{X}_1, \overline{X}_2, ..., \overline{X}_n, r)$ be obtained as follows. The graph consists of n nodes $v_1, v_2, ..., v_n$, such that v_i is located at $\overline{X}_i$. Two nodes v_i and v_j are connected by an edge if their distance from each other is less than r. Let $Con(\overline{X}_1, \overline{X}_2, ..., \overline{X}_n, r) = 1$ if the graph $g(X_1, X_2, ..., X_n, r)$ is connected and $Con(\overline{X}_1, \overline{X}_2, ..., \overline{X}_n, r) = 0$ otherwise. Then the probability that $g(n, r)$ is connected is exactly given by

$$\int_{(S_0)^n} Con(\overline{X}_1, \overline{X}_2, ..., \overline{X}_n) dm(\overline{X}_1, \overline{X}_2, ..., \overline{X}_n).$$

Although, it may not be very obvious, this formula is computationally infeasible. Thus, unless for very small values of n, such as $n = 2, 3$, it is practically useless. Moreover, this formula does not reveal anything about the interplay between different network parameters such as r and n and network properties. This example shows that obtaining exact formulas is not usually enough. Instead, we need to find meaningful and easily computable formulas.

Finally, we note that for some network quantities such as connectivity probability, it is possible to perform exhaustive simulations to estimate the quantity. Nevertheless, it is still very important to analytically study the network properties. First, analytic study helps us to understand the network behavior and see the effects of different parameters on the network properties. Thus, analytic results are very valuable in the design and evaluation of wireless networks. Second, there are many other network quantities that may not be evaluated by exhaustive simulations. For example, in this paper, we analytically study the capacity of wireless networks. It is not clear, if it is possible to set up simulations to estimate the network capacity. [1] Third, quantities such as connectivity probability are usually used in the analysis of more complicated network properties such as capacity analysis. Thus, it is important to analytically study them. Here is a simple analogy. In circuit design, we can always use the specialized computer packages to analyze a circuit. However, it is still very important to understand the behavior of different components of a circuit. A circuit designer must have access to analytic formulas and basic understanding of the circuit design methodology to design a circuit. Later computer simulations, can be helpful in validating the design, obtaining more exact evaluations, and making final adjustments.

V. FUNDAMENTALS OF SMALL-SCALE ANALYSIS

In this section, we try to establish a framework for analysis of finite networks. We list some important differences between small-scale and large-scale networks. In each subsection we first introduce the main idea, and then pick one or two network properties and show how to analyze those properties for small-scale networks. We try to choose simple examples that best show the difference between small-scale and large-scale analysis. In most cases, for simplicity, we only consider $g(n, r)$; while occasionally we give the results for the more general model $g(n, r, p)$. Nevertheless, it is not usually very difficult to extend the given results for $g(n, r)$, to $g(n, r, p)$.

A. Boundary Effects

One important phenomenon in asymptotic analysis is that boundary effects can be neglected. Loosely speaking, the anal-

[1] Note that we can estimate the average throughput for a given network with a specific protocol and data traffic model using exhaustive simulations. However, here by capacity we mean the highest possible achievable capacity, not the one achieved using a specific communication and routing protocol. Such capacity measure can be used to determine the efficiency of different protocols.

ysis of the network properties is usually dominated by what happens in region S_1 in Figure 1. In fact, we saw an example of this phenomenon in the asymptotic analysis of connectivity in Theorem 2. This can considerably simplify the analysis and results in simple and closed-form formulas for network properties. However, in small-scale networks boundary effects cannot be neglected. In other words, nodes in the corners of the field can play an important role in some of the network properties. To clarify this, let us consider a simple example. Suppose we want to find the average coverage in a wireless sensor network defined by $g(n, r)$. In other words, we want to find the average percentage area that is covered. For simplicity, suppose the sensing radius is also equal to r, that is, each node covers a circle of radius r centered at the node location. The probability that the point $\overline{X}$ in S_0 in not covered is given by

$$\left(1 - \nu(B(\overline{X}, r))\right)^n. \quad (3)$$

Thus, if PC_{notcov} is the average percentage of the uncovered area, then

$$PC_{notcov} = \int_{S_0} \left(1 - \nu(B(\overline{X}, r))\right)^n dm(\overline{X}). \quad (4)$$

Thus, PC_{notcov} can be obtained easily and the equation (4) is valid for all values of n, small and large. However, if we are doing asymptotic analysis, and we assume that $\lim_{n\to\infty} r(n) = 0$ (this assumption is almost always true for large-scale analysis), then using (4) we obtain

$$\begin{aligned} PC_{notcov} &= m(S_1)\left(1 - \pi r^2\right)^n + \\ &\int_{S_0 \setminus S_1} \left(1 - \nu(B(\overline{X}, r))\right)^n dm(\overline{X}) \\ &= (1 - o(1))(1 - \pi r^2)^n. \end{aligned} \quad (5)$$

Therefore, asymptotically, $PC_{notcov} = (1 - \pi r^2)^n$. Note that in this case the only difference between the exact (formula (4)) and asymptotic expressions comes from the edge effect. Figure 3 compares the two results. We observe that the two results differ considerably. This example clearly shows the importance of boundary effects in small-scale networks. This example is unique in the sense that the exact analysis is very simple. However, we are not usually so lucky. For instance, as we will see, exact analysis of other properties can be very complicated.

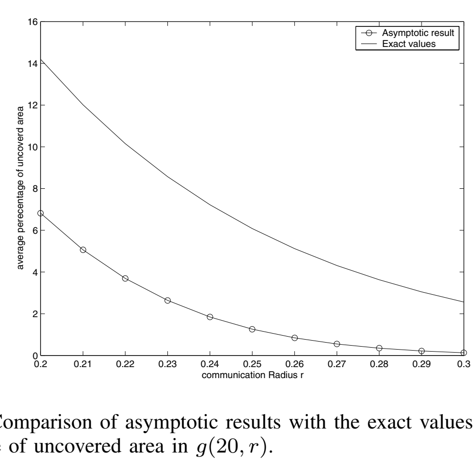

Fig. 3. Comparison of asymptotic results with the exact values for average percentage of uncovered area in $g(20, r)$.

Small-Scale Analysis for Connectivity Properties of g(n,r,p): Before discussing other differences between large-scale and small-scale analysis, we provide a small-scale analysis for connectivity properties of $g(n, r, p)$. This is a good example to illustrate our methodology for small-scale analysis. Since the exact analysis is usually very difficult or at least results in very complicated formulas, a good approach is to find simple lower and upper bounds. Therefore, in this section we find lower and upper bounds for the probability that $g(n, r, p)$ is disconnected, $p_{disc}(n, r, p)$. As we will see the two bounds almost coincide with each other. Thus, they give a very good estimate for $p_{disc}(n, r, p)$. Indeed, the two bounds completely agree with the simulation results. Let $p_{low}(n, r, p)$ and $p_{upp}(n, r, p)$ be the lower and upper bounds on $p_{disc}(n, r, p)$, respectively. Here we consider the case where $p_{disc}(n, r, p)$ is small, i.e., $p_{disc}(n, r, p) < .1$. In practice, this is usually the range that is important, since we want that the network is connected with high enough probability.

Theorem 3: Consider a wireless network with the model $g(n, r, p)$. Then we have

$$\begin{aligned} p_{disc}(n, r, p) \;&\geq n \int_{S_0} \left(1 - \nu(B(\overline{X}, r))p\right)^{n-1} dm(\overline{X}) - \\ &\binom{n}{2} \int_{S_0}\int_{S_0} \Big(1 - \nu(B(\overline{X}, r))p - \nu(B(\overline{X}, r))p + \\ &\nu(B(\overline{X}, r) \;\cap B(\overline{Y}, r))p^2\Big)^{n-2} dm(\overline{X}) \times m(\overline{Y}). \end{aligned} \quad (6)$$

Proof: For proof, please see the longer version of the paper available at http://www.ecs.umass.edu/ece/pishro/. ■

Note that this lower bound for $p_{disc}(n, r, p)$ may seem to be too complicated and thus may not satisfy the simplicity requirement. However, as we will see, this lower bound is almost the same as a simple upper bound that we find shortly. Thus, the simple upper bound can be used in estimating $p_{disc}(n, r, p)$. The lower bound is useful in the sense that it assures us that our estimate is very close to the real value for $p_{disc}(n, r, p)$.

We now find an upper bound for $p_{disc}(n, r, p)$. By definition, a connected component of a graph g is a connected subgraph that is isolated from the rest of g. Thus, $p_{disc}(n, r, p)$ is equal to the probability that $g(n, r, p)$ has at least one component of size less than $n/2$. For $U \subseteq \{v_1, v_2, ..., v_n\}$, let $p_{comp}(U)$ be the probability that the vertices in U construct a connected component in $g(n, r, p)$. Then, we have

$$p_{disc}(n, r, p) \leq \sum_{k=1}^{n/2} \binom{n}{k} p_{comp}(\{v_1, v_2, ..., v_k\}) = \sum_{k=1}^{n/2} a_k. \quad (7)$$

Note that although $p_{upp}(n, r, p) = \sum_{k=1}^{n/2} a_k$ is a valid upper bound for $p_{disc}(n, r, p)$, it does not satisfy the simplicity requirement. In fact, except the first few terms, computing a_k's is computationally infeasible. We now try to simplify this upper bound. Note that so far all the results about the lower and upper bounds have been exact and rigorous. However, we now use a simple approximation for simplifying the upper

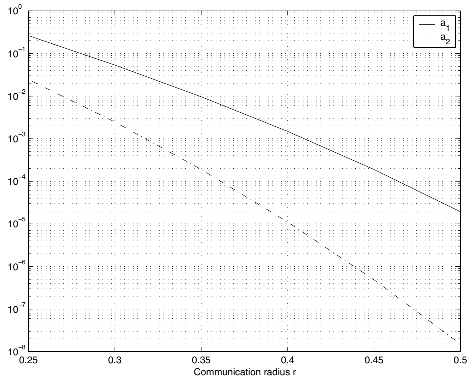

Fig. 4. Comparison of a_1 and a_2 in (7).

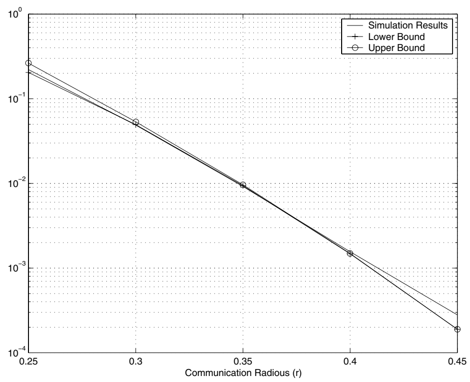

Fig. 5. Disconnectivity probability of $g(100, r, .5)$: lower bound, upper bound, and the simulation results.

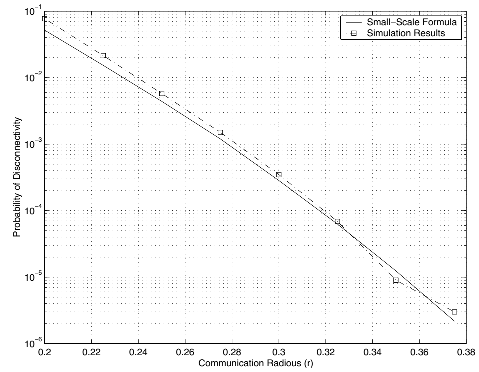

Fig. 6. Disconnectivity probability of $g(100, r, 1)$ using (10) and simulation results.

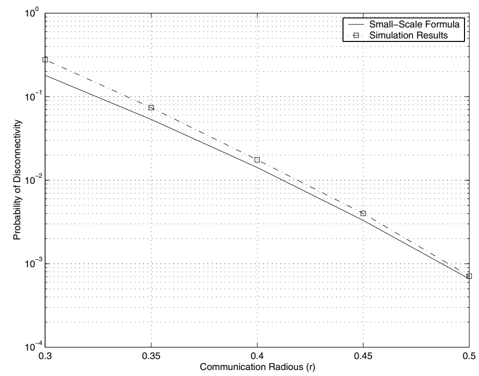

Fig. 7. Disconnectivity probability of $g(30, r, 1)$ using (10) and simulation results.

bound. Nevertheless, the approximation is completely backed by numerical and analytical arguments. We remember our assumption that $p_{disc}(n, r, p)$ is not very large, specifically we assume $p_{disc}(n, r, p) < .1$. An important observation here is that, by this assumption, the a_k's decay very fast and the $p_{upp}(n, r, p) = \sum_{k=1}^{n/2} a_k$, is dominated by a_1. This can be seen by both numerical simulations and intuitive analytical arguments. To see this let us examine a_1 and a_2. Figure 4 compares a_1 and a_2 for $g(n = 100, r, p = .5)$. As we see a_2 is at least one order of magnitude lower than a_1. Note that this is a crucial observation that simplifies the upper bound significantly.

The fact that a_k's decay very fast, can also be described in the following way. For a subset of vertices $U = \{u_1, u_2, ..., u_t\} \subseteq \{v_1, v_2, ..., v_n\}$, let $A(U)$ be the area of the unions of circles with radii r centered at u_i's. Then the probability that the vertices in U are isolated from the rest of the graph is given by

$$\big(1 - A(u)\big)^{n-t} \simeq e^{-nA(u)}. \tag{8}$$

This shows that $p_{comp}(\{v_1, v_2, ..., v_k\})$ in (7), decays exponentially fast with the number of vertices, k. Thus, a_k's decay very fast. This is of course consistent with our observation in Fig. 4. Therefore, we conclude

$$p_{upp}(n, r, p) \simeq a_1 = n \int\limits_{S_0} \left(1 - \nu(B(\overline{X}, r))p\right)^{n-1} dm(\overline{X}). \tag{9}$$

Figure 5 shows the upper bound, lower bound, and the simulation result for the probability of disconnectivity of $g(n, r, p)$, for $n = 100$, and $p = .5$. As we see the three almost coincide. As we will see shortly, similar results are achieved if we use different choices of parameters. Thus, we conclude

$$p_{disc}(n, r, p) \simeq n \int\limits_{S_0} \left(1 - \nu(B(\overline{X}, r))p\right)^{n-1} dm(\overline{X}). \tag{10}$$

Note that (10) suggests that $p_{disc}(n, r, p)$ is dominated by the probability of having an isolated vertex. We recall from Theorem 2 that the asymptotic probability of disconnectivity, $1 - p_{asymp}(c)$, is also dominated by the isolated vertices. However, a crucial difference between these two is that the boundary effects are insignificant in asymptotic analysis. This causes that the asymptotic formula differs from the correct values by several orders of magnitude when used for small or moderate values of n as shown in Fig. 2. However, our small-scale formula is almost identical to the correct values because it considers the boundary effects. Note that (10) gives us a very simple and easily computable formula for disconnectivity probability.

Figures 6, 7, and 8 compares the disconnectivity probabilities obtained by (10) and simulations for different values of n and p. We confirm that in all the cases the given formula matches the simulation results.

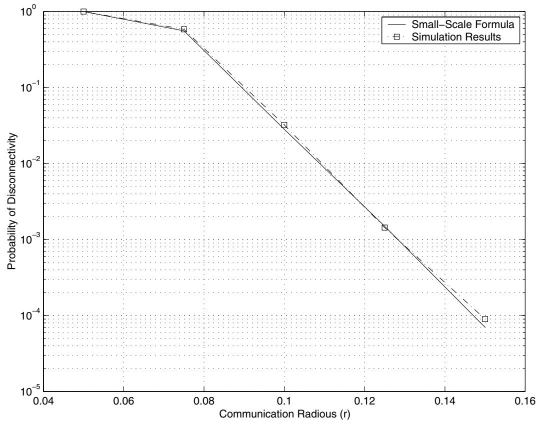

Fig. 8. Disconnectivity probability of $g(500, r, 1)$ using (10) and simulation results.

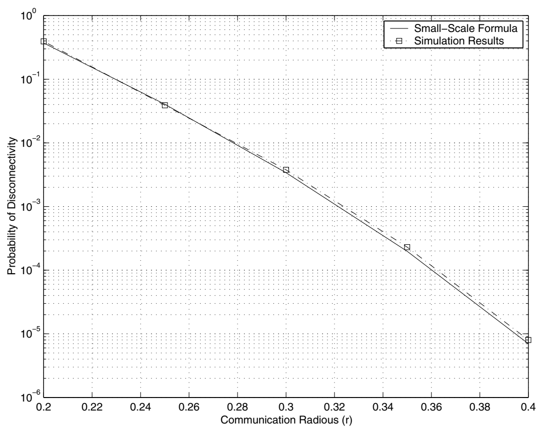

Fig. 9. Probability that $g(100, r, 1)$ is not two-connected, using (10) and simulation results.

It is worth noting that the methodology used here can be used to study k-connectivity which is more general than connectivity. As it was mentioned earlier, k-connectivity is important for multi-path routing, reliability, and security in networks. By definition, a network is k-connected if there does not exist a set of $k-1$ vertices whose removal disconnects the graph. In particular, 1-connectivity ($k = 1$) is equivalent to connectivity. Using similar arguments, we find that the probability that $g(n, r)$ is not k-connected, $p_{k,disc}(n, r)$ is dominated by the probability that there exists at least one vertex in the network with degree less than k. In summary, we find the following approximation of the probability that $g(n, r)$ is not k-connected

$$p_{k,disc}(n,r) \simeq \sum_{j=0}^{k-1} n \binom{n}{j} \int_{S_0} [\nu(B(\overline{X}, r(n)))]^j \times \left(1 - \nu(B(\overline{X}, r(n)))\right)^{n-j-1} dm(\overline{X}). \quad (11)$$

Figure 9, validates this expression for $k = 2$. Again we verify that the formula matches the simulation results. Our simulations for larger values of k consistently confirms the validity of (11). Here, due to the space limitations we omit these results.

B. Effect of Constant Factors

So far, we have seen the importance of boundary effects in the analysis of finite networks. We now discuss another important issue. In asymptotic analysis, we usually neglect constant factors. However, in small-scale analysis, we must consider them. This is in fact, a difference between any finite analysis and asymptotic analysis and is not specific to geometric graphs. To show the importance of constant factors in the geometric graphs of wireless networks, we consider the medium access (MAC) layer capacity. Asymptotic MAC-layer capacity of ad hoc wireless networks is studied in [1]. The MAC-layer capacity is defined in [1] as the maximum possible number of concurrent transmissions at the media access layer. It is shown in [1] that for a wide class of MAC protocols including IEEE 802.11, the MAC-layer capacity can be modeled as a maximum Distance-2 matching (D2EMIS) problem in the underlying wireless network. That is, given a graph $G(V, E)$, find a maximum set of edges $E' \subseteq E$ such that no two edges in E' are connected by another edge in E. It is shown in [1] that for $g(n, r)$, the MAC-layer capacity is optimized at $r = \Theta(\frac{1}{\sqrt{n}})$ and is given by $\Theta(n)$. Although this is an important and valuable result, it has very limited value when we consider finite networks. For example, suppose we have a network consisting of 100 sensors and we want to choose the communication radius such that the MAC-layer capacity is optimized. The asymptotic result does not tell us what the value of r should be. Moreover, we do not know what the optimum MAC-layer capacity would be. This example clearly shows the importance of constant factors in small-scale analysis. In the next section we analyze the average MAC-layer capacity for finite networks and obtain simple lower and upper bounds. Using these bounds, we try to answer the above question about the MAC-layer capacity of a finite sensor network.

Small-Scale Analysis of MAC-Layer Capacity of $g(n, r)$: In this section we analyze the average MAC-layer capacity of $g(n, r)$, i.e the maximum number of possible concurrent transmissions which is available on average in $g(n, r)$. As it was mentioned, we find simple upper and lower bounds and by which we find the optimum value of r and the corresponding average MAC-layer capacity. We now prove the following lower bound. Let $MAC(n, r)$ be the average MAC-layer capacity of $g(n, r)$.

Theorem 4: Define

$$s = \int_{S_0} \nu(B(\overline{X}, 2r)) dm(\overline{X}), \quad (12)$$

$$t = \frac{1 - (1 - s)^n}{s}. \quad (13)$$

Then, the average MAC-layer capacity satisfies

$$MAC(n, r) \geq t\left(1 - (1 - \pi r^2)^{n-t}\right). \quad (14)$$

Proof: For proof, please refer to the longer version of the paper available at http://www.ecs.umass.edu/ece/pishro/. The proof is constructive. That is, we use an algorithm to find a set of m concurrent transmissions in $g(n, r)$ such that on average m satisfies the lower bound given by the theorem. ■

We now obtain a simple upper bound on the average MAC-layer capacity.

Theorem 5: Consider a wireless network graph $g(n, r)$. Define

$$n_1 = n \int_{S_0} (1 - \nu(B(\overline{X}, r)))^{n-1} dm(\overline{X}). \quad (15)$$

Then, the average MAC-layer capacity satisfies

$$MAC(n, r) \leq \frac{n - n_1}{2 + 1.37 r^2 n}. \quad (16)$$

Proof: For proof, please see the longer version of the paper available at http://www.ecs.umass.edu/ece/pishro/. ■

Figure 10 shows the upper and lower bounds on the MAC-layer capacity of $g(100, r)$. The lower bound is maximized at $r = .06$, while the upper bound is maximized at $r = .08$. Thus, to optimize the MAC-layer capacity we can choose $.06 < r < .08$. We also note that the maximum achievable MAC-layer capacity is between 15 and 30. An interesting open problem is to tighten the bounds to obtain a more accurate estimate of MAC-layer capacity.

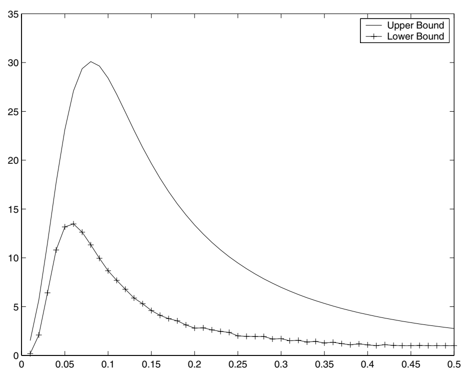

Fig. 10. Upper and lower bounds on the average MAC-layer capacity of $g(100, r)$.

C. Lack of Concentration

In asymptotic analyses, usually random variables concentrate on their average values. Thus, it usually suffices to only determine the expected value. However, in small-scale analysis this is not the case. Thus, knowing the expected value is not usually enough. To clarify this, it is useful to consider geometric routing algorithms. A survey on routing protocols for wireless sensor networks can be found in [31]. Suppose nodes A and B are two fixed nodes on the plane that are located at the unit distance away from each other. In geometric routing when node A wants to send the data to node B, the information is usually sent hop by hop to B. Each node passes the data to another node which is somewhat closer to the destination, node B. An example of such routing algorithms is GPSR [32]. Since in these algorithms the next hop is typically determined locally, when the density of nodes is large, by a martingale argument we can usually prove that the path length (the number of hops) is concentrated around its average with high probability. Thus, determining the average path length would be sufficient. On the other hand, for small-scale networks, the number of hops can considerably deviate from the average. Thus, analysis of geometric routing protocols for small-scale networks can be very important. To show this, we now analyze a simple geometric routing algorithm. We should note that the algorithm is not optimal in terms of energy, delay, etc. However, it is good enough to show the distinct characteristics of small-scale networks.

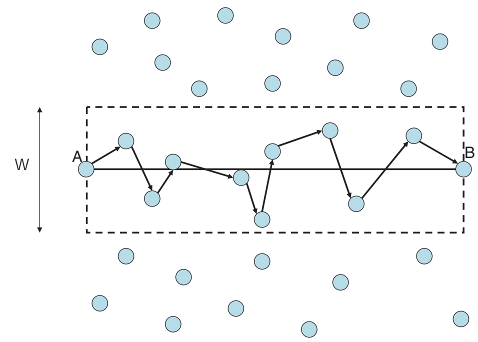

Fig. 11. Illustration of a simple geometric routing.

Consider the following scenario. As shown in Figure 11, let us assume that node A wants to send some information to node B. Suppose the nodes are distributed on the plane based on a Poisson distribution with density λ. Assume the distance between A and B is $d(A, B) = 1$. We connect A and B by a virtual line and also consider a virtual rectangle shown in Fig. 11 with width w. The routing path consists of all the nodes in the rectangle, from left to right. In this routing scenario, we assume that the packets travel from a node to its right neighbor node with the shortest horizontal distance. Assume that A is located at $(0, 0)$, and B at $(1, 0)$, and the ith node in the rout is located at (X_i, Y_i). Then, $X_{i+1} - X_i$ has an exponential distribution with parameter λw. Thus, if H is the number of hops from A to B, we have

$$\text{Prob}\{H = h\} = \text{Gamcdf}\,(1, h-1, \frac{1}{w\lambda}) - \text{Gamcdf}(1, h, \frac{1}{w\lambda}),$$

where Gamcdf(x, h, η) is the value of the Gamma distribution function with parameters h and η at point x. By considering the Gamma distribution, we can show that for small $\eta = \lambda w$, the distribution is very wide (refer to the longer version of the paper available at http://www.ecs.umass.edu/ece/pishro/). However for larger η, the distribution concentrates around its average, $EH = \eta + 1$. This shows that although in the asymptotic case the average value can suffice for the analysis, the whole distribution should be known in the finite density.

As an application of this, let us consider the energy issue. In wireless sensor networks, energy is arguably the most important constraint. Thus, we would like to minimize the energy consumption. We assume that the energy needed for a direct transmission from a node to a neighbor at distance d is proportional to d^2. Here, we assume that every sensor adjusts its transmission power according to its distance from the recipient node. Then, using the distribution of the random variables involved in our simple geometric routing, we conclude that the average total energy consumption in communication between A and B is proportional to

$$Avg.Energy \propto (\lambda w + 1)\left(\frac{2}{\lambda^2 w^2} + \frac{w^2}{12}\right). \quad (17)$$

Thus, for a given λ, we can find the value of w that minimizes the average energy. It is very important to note that the geometric algorithm used here is not the best possible, and the assumption of the energy adjustment in the sender node may not be realistic in some scenarios. However, almost all geometric routings have similar properties. Specifically, in all geometric algorithms, when the density of nodes tends to infinity, the number of hops converges to the average value, while when the network is not very dense the number of hops can deviate considerably.

VI. Conclusion

In this paper, we provided a methodology for analysis of wireless networks in finite regime. We provided some compelling evidence to show that asymptotic results are not suitable for analyzing practical finite networks. We considered connectivity, coverage, MAC-layer capacity and routing algorithms of finite networks. We obtained a very simple formula for connectivity of wireless networks and verified it by simulation results. The formula was then extended to include k-connectivity. We studied MAC-layer capacity and obtained simple lower and upper bounds. Using these bounds we estimated the optimum value for achieving the highest capacity. Finally, we studied geometric routings. Using, these examples we confirmed that finite-scale networks posses unique characteristics that require a new framework distinct from asymptotic approaches.

This paper opens up many research possibilities that offer potential for further research. In the past, many other important properties of wireless networks have been studied for large-scale networks. It is an important task to extend these results for networks with practical sizes, i.e., small-scale networks. For example, there are several other measures for network capacity such as transport capacity, information theoretic capacity, and capacity of cooperative nodes. Asymptotic analysis of these definitions has been studied extensively. It is very useful to extend these results to small-scale networks. Finite analysis can reveal the effects of network parameters on networks characteristics.

References

[1] H. Balakrishan, C. L. Barrett, V. S. A. Kumar, M. V. Marathe, and S. Thite, "The distance-2 matching problem and its relationship to MAC-layer capacity of ad hoc wireless networks," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1069 –1079, 2004.

[2] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[3] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad-hoc wireless networks," *IEEE Infocom, Anchorage, Alaska, USA*, April 2001.

[4] P. Gupta and P. R. Kumar, "Towards an information theory of large networks: an achievable rate region," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1877–1894, 2003.

[5] J. Li, C. Blake, D. S. J. De Couto, H. I. Lee, and R. Morris, "Capacity of ad hoc wireless networks," in *Proceedings of the 7th ACM International Conference on Mobile Computing and Networking*, (Rome, Italy), pp. 61–69, July 2001.

[6] E. Perevalov and R. Blum, "Delay limited capacity of ad hoc networks: Asymptotically optimal transmission and relaying strategydelay limited capacity of ad hoc networks: Asymptotically optimal transmission and relaying strategy," *IEEE Infocom, San Francisco, CA, USA*, 2003.

[7] B. Liu, Z. Liu, and D. Towsley, "On the capacity of hybrid wireless networks," *IEEE Infocom, San Francisco, CA, USA*, 2003.

[8] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, pp. 102–114, August 2002.

[9] B. Bollobás, *Random Graphs*. Cambridge University Press, second ed., 2001.

[10] R. Meester and R. Roy, *Continuum Percolation*. Cambridge University Press, cambridge UK, 1996.

[11] M. Penrose, *Random Geometric Graphs*. Oxford University Press, 2003.

[12] P. Gupta and P. Kumar, "Critical power for asymptotic connectivity in wireless networks," *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming, W.M. McEneaney, G. Yin and Q. Zhang (Eds.)*, 1998.

[13] F. Xue and P. R. Kumar, "The number of neighbors needed for connectivity of wireless networks," *Wireless Networks*, vol. 10, no. 2, pp. 169–181, 2004.

[14] L. Booth, J. Bruck, M. Franceschetti, and R. Meester, "Covering algorithms, continuum percolation and the geometry of wireless networks," *Annals of Applied Probability*, vol. 13, May 2003.

[15] L. Booth, J. Bruck, M.Cook, M. Franceschetti, and R. Meester, "Continuum percolation with unreliable and spread out connections.," submitted.

[16] S. Shakkottai, R. Srikant, and N. Shroff, "Unreliable sensor grids: Coverage, connectivity and diameter," *In the proceedings of IEEE INFOCOM'03, San Francisco, CA, April 2003*.

[17] D. Dubhashi, O. Hggstrm, and A. Panconesi, "Connectivity properties of bluetooth wireless networks," submitted.

[18] X. Y. Li, P. Wan, Y. Wang, and C. W. Yi, "Fault tolerant deployment and topology control in wireless networks," *ACM Symposium on Mobile Ad Hoc Networking and Computing, MOBIHOC 2003*.

[19] P. Wan and C. W. Yi, "Asymptotic critical transmission radius and critical neighbor number for k-connectivity in wireless ad hoc networks," *ACM Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc 2004*.

[20] S. Kumar, T. H. Lai, and J. Balogh, "On k-coverage in mostly sleeping sensor network," in *MobiCom*, September 2004.

[21] H. Pishro-Nik, K. Chan, and F. Fekri, "On connectivity properties of large-scale sensor networks," *First Annual IEEE International Conference on Sensor and Ad Hoc Communications and Networks*.

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

[23] O. Dousse and P. Thiran, "Connectivity vs capacity in dense ad hoc networks," *IEEE Infocom*, 2004.

[24] M. Franceschetti and R. Meester, "Critical node lifetimes in random networks via the chen-stein method," Available online at http://fleece.ucsd.edu/ massimo/.

[25] M. Desai and D. Manjunath, "On the connectivity in finite ad hoc networks," *Communications Letters, IEEE*, vol. 6, pp. 437–439, October 2002.

[26] A. Gore, "Comments on the connectivity in finite ad hoc networks," *Communications Letters, IEEE*, vol. 10, pp. 88–90, February 2006.

[27] N. Karmachandani, D. Manjunath, D. Yogeshwaran, and S. K. Iyer, "Evolving random geometric graph models for mobile wireless networks," in *Proceedings of the Fourth International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, (WIOPT 2006)*, 2006.

[28] L.-H. Yen and C. W. Yu, "Link probability, network coverage, and related properties of wireless ad hoc networks," in *Proceedings of the First International Conference on Mobile Ad-Hoc and Sensor Systems (MASS 2004)*, October 24–27 2004.

[29] A. Ghasemi and S. Nader-Esfahani, "Exact probability of connectivity in one-dimensional ad hoc wireless networks," *IEEE Communications Letters*, vol. 10, pp. 251–253, 2006.

[30] K. C. H. Pishro-Nik and F. Fekri, "On connectivity properties of large-scale wireless sensor networks," in *Proc. First Annual Communications Society Conf. on Sensor Communications and Networks Santa Clara*, pp. 498–507, IEEE, 4-7 October 2004.

[31] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Elsevier Ad Hoc Network Journal*. To appear.

[32] B. Karp and H. T. Kung, "GPSR: gready primeter stateless routing for wireless sensor networks," in *The 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking(MobiCom '00)*, 2000.