Critical Sensor Density for Partial Connectivity in Large Area Wireless Sensor Networks

HAIYAN CAI
University of Missouri-St. Louis
XIAOHUA JIA
City University of Hong Kong
and
MO SHA
Washington University in St. Louis

In this paper, we study the critical sensor density for partial connectivity of a large area sensor network. We assume that sensor deployment follows the Poisson distribution. For a given partial connectivity requirement ρ, $0.5 < \rho < 1$, we prove that there exists a critical sensor density λ_0, around which the probability that at least a fraction $\rho\%$ of sensors are connected in the network increases sharply from ε to $1 - \varepsilon$ within a short interval of sensor density λ. The length of this interval is in the order of $O(-\log \varepsilon / \log A)$, where A is the area of the sensor field, and the location of λ_0 is at the sensor density where the above probability is about 1/2. We prove the above theoretical results in the hexagonal model. We also extend our results to the disk model that models transmission range of sensors as disks. Simulation results have verified our theoretical results and exhibited a close match of the results in the hexagon model and the disk model.

Categories and Subject Descriptors: C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*network topology*; F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Nonnumerical Algorithms and Problems—*Geometrical problems and computations*

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Wireless Sensor Networks, critical density, simulation.

1. INTRODUCTION

The problem of connectivity of sensors deployed randomly in a large area is a central issue in the studies of such networks. Each sensor connects only locally to the sensors within its communication range. The connections among the sensors

Part of this article was published in Proceedings of the Infocom [Cai et al. 2010]. This work was supported in part by a grant from NSF China Grant No. 60633020 and 60970117.
Author's address: Haiyan Cai, Department of Mathematics and Computer Science,University of Missouri-St. Louis, St. Louis, MO. 63121.
Xiahua Jia, Department of Computer Science,City University of Hong Kong, Hong Kong.
Mo Sha, Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, MO. 63130.

across the field rely on intermediate relays. Extensive studies have been done on connectivity of sensor networks. Some early works focused on how many neighbors each node should connect to in order to maximize the network throughput, such as [Kleinrock and Silvester 1978], [Takagi and Kleinrock 1984], [Hou and Li 1986]. More recent works studied how large the node transmission range is needed to ensure the full connectivity of the network (with high probability), or equivalently how many nodes are needed to ensure such connectivity given node transmission range is fixed [Bettstetter 2002], [Wan and Yi 2004], [Santi and Blough 2003]. Some other works investigated the relationship between network connectivity and sensor coverage and studied the deployment of minimal number of sensors in a field such that the network is connected and the field is fully covered [Xing et al. 2005], [Zhang and Hou 2005], [Bai et al. 2006].

However, most of the above studies focus only on the full connectivity (or k-connectivity) of the sensors in which every single sensor is connected to the entire network. This is a rather stringent requirement. When the sensor deployment is random, the requirement of full connectivity inevitably creates excessive redundant sensors, in order to guarantee the connections of a few otherwise isolated sensors. Although the price of sensor devices keeps decreasing in recent years, considering the large scale deployment, they can still be quite expensive (usually around \$100/piece). To address the issue, we investigate the problem of partial connectivity of the sensors in this paper.

For a given percentage requirement ρ, we are interested in the events like "the percentage of the sensors that are connected is at least ρ", or "the percentage of the area covered by some connected sensors exceeds ρ". We want to find how, in general, the probabilities of these events are related to the sensor density. It is clear that a higher density will always increase the chance of occurrence of such an event, and the question of our concern is, for a fixed sensor field, how fast such a chance will increase when the sensor density increases. Through our analysis, we find that there exists a critical value of the sensor density λ_0 such that the probabilities of the above events increase sharply from some small $\varepsilon > 0$ to $1-\varepsilon$ in a short interval centered near λ_0. Our analysis also reveals that, roughly speaking, the length of this interval is of the order $O(-\log\varepsilon/\log A)$, where A is the area of the field. In other words, if the probability for partial connectivity is ε when the sensor density if λ' and is $1-\varepsilon$ when the sensor density is λ'', then $\lambda''-\lambda' = O(-\log\varepsilon/\log A)$ for large A. We further find that λ_0 is such a value at which the probability for the event to occur is about $1/2$ when A is large. More precise statements on these results are given in later sections of this paper.

The results have the following implications. For random sensor deployment in a large field, we can always determine a critical value of the sensor density for a required percentage of node connectivity (or area coverage) such that when the density is slightly smaller than this critical value, the probability of meeting the connectivity percentage is very low, and when the density is increased to slightly above this critical value, the probability of meeting the requirement will be very high. Moreover, this transition of the probability from a very small value to a value nearly equal to 1 occurs in an interval whose length goes to 0 as the area of the sensor field goes to infinity. Note that, in this study we consider a more general requirement of connectivity (or coverage), i.e., partial connectivity, unlike most of the existing works that assumed full (100%) connectivity requirement. As the simulation study will show below, the full connectivity (or coverage) would require much greater number of sensors in the field just to avoid the exclusion of a small fraction of isolated sensors. The savings from reducing the number of sensors will be significant if we are willing to give up a small portion of the sensor connectivity.

We will establish the results with the following approach. We first divide the sensor field into a hexagonal lattice and obtain the analytical results of sensor connectivity against sensor density in the hexagon model. Then, we extend our discussion to the usual disk model for modeling both sensor communication and sensing range. Finally, we verify the theoretical results by comparing the hexagonal model with the disk model through extensive simulations, and we seek the match between the two models by adjusting the hexagon size according to the sensor transmission radius.

2. RELATED WORK

Connectivity of the network highly depends on node density. Node density can be controlled by deploying more nodes in a field, or equivalently by increasing node transmission radius. Early studies of network connectivity focused on achieving high network throughput. This is because high connectivity would cause high interference, and thus lead to poor throughput. Kleinrock and Silvester proposed in [Kleinrock and Silvester 1978] that for maximizing the one-hop progress of packets towards desired receivers under the slotted ALOHA protocol, every node on average should connect itself to six nearest neighbors. This magic number "six" was changed to five and seven for some other transmission protocols in [Takagi and Kleinrock 1984]. Hou and Li presented a model in [Hou and Li 1986] to consider scenario that each node can adjust its transmission radius individually and obtained the result that each node should on average connect to 6∼8 neighbors in order to maximize the network throughput. However, these works did not consider the connectivity of the whole network. As pointed out by Ni and Chandler in [Ni and Chandler 1994], by connecting 6∼8 nearest neighbors, a small size network would have a high probability of being connected, but it is almost certain that a large size network would be disconnected. In fact, Gilbert [Gilbert 1961] found that there was a critical number N_0, $1.64 < N_0 < 17.9$, such that if every node connects to N_0 nearest neighbors, the random plane network contains an infinite connected component with nonzero probability. The range of N_0 was later tightened to $2.195 < N_0 < 10.526$ by Philips et al. in [Philips et al. 1989]. Recently, Xue and Kumar proved in [Xue and Kumar 2004] that if each node connects to less than $0.074 \log n$ nearest neighbors, the probability that the network is connected converges to 0 as the increase of n (n is the number of nodes in the networks); and if each node connects to greater than $5.1774 \log n$ nearest neighbors, the network is asymptotically connected.

Some other works studied the problem of how large the node transmission radius is needed to ensure the connectivity of the networks (equivalently, given node transmission radius, how many nodes are needed in a certain area to ensure the connectivity). Bettstetter [Bettstetter 2002] derived an analytical expression, for a given network density and a random uniform distribution of nodes, to find out the required transmission radius such that the induced network is almost surely (i.e., with very high probability) k-connected. Wan and Yi [Wan and Yi 2004] studied the asymptotic critical transmission radius for nodes with uniform radius and the asymptotic critical neighbor number for nodes that each of them can adjust their transmission radius individually, to ensure the induced network is k-connected. They provided a precise asymptotic distribution of transmission radius and an improved asymptotic upper bound on the critical neighbor number for k-connectivity. Santi et al. [Santi and Blough 2003] [Santi et al. 2001] studied the similar problem that given a set of nodes randomly and uniformly distributed in a d-dimensional region (d = 1, 2, and 3) with a fixed side length, to find the node transmission radius to ensure the resulting network is connected with high probability. They derived upper and lower bounds on the critical radius for one-dimensional networks, and non-tight bounds for two and three-dimensional networks. Through the simulations, they also studied the minimum transmission radius that ensures a large connected component of the network (not fully connected). The above works all studied the asymptotic density to ensure the network connectivity with high probability. There is no theoretical proof on the existence of a critical density λ_0 at which there is a sharp increase of connectivity for the general case of partial connectivity and no theoretical analysis for finding the exact value of such λ_0. An original contribution of this paper is to provide theoretical analysis on the critical density for connectivity and to find the exact value of this critical density.

In wireless sensor networks, network connectivity problem is closely related to the sensing coverage problem. This relationship between connectivity and coverage was pointed out early by Piret in [Piret 1991] and he gave a conjecture stating that the transmission radius of a set of nodes in an area must be at least twice as much as their sensing radius to ensure the set of nodes connected while their sensing range covers the whole area. More recently, Xing et al. [Xing et al. 2005] presented a geometric analysis of the relationship between connectivity and coverage, and proposed a coverage configuration protocol that can provide different degrees of coverage (i.e., k-coverage) for applications. Zhang and Hou also addressed this problem in [Zhang and Hou 2005] and proposed a decentralized density control protocol, which selects a minimum subset of densely deployed nodes that are connected and can cover the entire area. Bai et al. [Bai et al. 2006] considered the problem for placing the minimum number of sensors in a rectangular region such that the sensor network is one or two-connected and the entire region is covered. They proposed an optimal deployment pattern. Their work was later extended to a more general case of finding the optimal deployment that achieves k-connectivity (k=1, 4, 6) and full-coverage in [Bai et al. 2009].

Although most studies focus on full connectivity of the sensor networks, the concept of partial connectivity is not new [Dousse et al. 2006]. Indeed, in the context of continuum percolation [Meester and Roy 1996], the connectivity problem of the sensor network can be formulated using a percolation model in Poisson blobs. In this model, when the sensor density exceeds certain critical threshold, a percolation occurs, meaning that with a non-zero probability $\theta(\lambda) > 0$ there exists an infinite network of connected sensors in an infinite sensor field. Furthermore, as it was shown in [Penrose and Pisztora 1996], for any given sensor density above the percolation threshold, by restricting the sensor field to a sufficiently large but finite sub-field, one can achieve a percentage of connectivity with a probability approaching to 1 exponentially fast, as the area of the sub-field goes to infinity. The percentage of the connectivity converges to $\theta(\lambda)$. This result is, however, different from ours. The issue we are exploring here is how fast the probability of a partial connectivity with a given connection percentage will increase, when the sensor area is large but *fixed*, and the sensor density is allow to change. More precisely, our problem is, given a sensor field *with fixed area*, where to locate and how to estimate the range of the sensor density within which the probability will change from nearly 0 to nearly 1. This later question can be more important in a real sensor deployment situation where a sensor area is pre-fixed and one needs to decide how many sensors to use. In this paper we provide an answer to the later question by establishing a sharp-threshold property around a critical value of the sensor density. Note that the critical value we introduce here for the partial connectivity is different from the critical value for a percolation. Also see [Osher 2007] for other relevant discussions.

3. PROPERTIES OF CONNECTIVITY PROBABILITIES

We now formally describe our problems and the results. Let λ be the sensor density, the number of sensors per unit area. Suppose N_A sensors are deployed randomly and independently in a square region of area A, where we assume N_A is a Poisson random variable $N_A \sim Poisson(\lambda A)$. Note that this is equivalent to assuming that the sensors are located according to a Poisson point process. In sensor networks, the communication range of a sensor is usually modeled as a disk with transmission radius R_c. Two sensor nodes can communicate with each other if they are in each other's communication range. We are interested in the probability that certain percentage of sensors are connected under a given density λ. A direct analysis of this probability is difficult. We are, however, able to study the probability in a slightly simplified model, i.e., the hexagon model. In what follows, we will first formulate the hexagonal model and study the connectivity problem of the sensor network in this model. We will then use the results from the hexagon model to approximate the connectivity probability in the disk model.

3.1 The Hexagon Model

We partition the sensor field into hexagons using a regular honeycomb hexagon lattice. Hexagons are arranged in such a way that each hexagon has two of its six sides placed vertically. To have a fixed orientation, we assume hexagons with adjacent vertical sides line up to form a row. Let M be a positive integer. Suppose that in the lattice there are M rows and that in each row, there are M hexagons. Let $H_{i,j}$ denote the jth hexagon in the ith row, $i,\ j\ = 1, 2, \cdots, M$. Every off-boundary hexagon $H_{i,j}$ has six neighbors, including $H_{i,j-1}$ and $H_{i,j+1}$ on each side. Let $H = \{H_{i,j} :\ i,\ j\ = 1, 2, \cdots, M\}$. Note that for large M, $A = \Theta(M^2)$. Of course, other regular lattices can also be considered for the modeling purpose. The reason we choose hexagon lattice is that some of its properties will be needed in the later proofs.

Let A_H be the area of the hexagons. From the previous assumption, there is an equal probability for a sensor to fall into each $H_{i,j}$. The probability that a hexagon will contain at least one sensor depends on λ. We denote this probability as $p(\lambda)$. Then

$$p(\lambda) = 1 - e^{-\lambda A_H}. \tag{1}$$

We say a hexagon is occupied, if there is at least one sensor in it. Thus, $p(\lambda)$ is the probability that a hexagon is occupied when the sensor density is λ. We say two neighboring hexagons are directly-connected if they are both occupied. Two hexagons in the lattice are said to be connected, if there is a path of directly-connected hexagons that connects them. Hexagons that are connected with each other form a cluster (a maximal connected component) in the corresponding graph. In this model, sensors are divided into a set of disconnected clusters.

In the hexagon model, we assume that two sensors communicate with each other if and only if they are either in the same hexagon or in the neighboring hexagons. By adjusting the hexagon size we can always obtain either a stronger or a weaker connectivity assumption in the hexagon model. Such a flexibility will allow us to obtain bounds for the probabilities of the disk model. For example, to obtain a lower bound, we can assume a stronger connectivity hexagon model by setting the hexagons size so that the farthest distance between two points inside two neighboring hexagons is R_c, the communication range of the disk model. That is, the hexagon edge should have a length $R_c/\sqrt{13}$ (see details in Figure 2(a)).

According to this definition of connectivity, two sensors are connected if and only if they are in the same cluster of hexagons. We are interested in the percentage of sensors which are connected within a given cluster. Let C be a cluster in the network, and N_C the total number of sensors in C. Let us consider the ratio

$$r_N(C) = \frac{N_C}{N_A}. \tag{2}$$

For a given ρ, $1/2 < \rho < 1$, we defined the event

$$B_\rho = \{\text{There is a cluster } C \text{ such that } r_N(C) \geq \rho\}. \tag{3}$$

B_ρ is the event that at least a fraction $\rho\%$ of the sensors in the sensor field are connected. Note that since $\rho > 1/2$, there can be at most one such cluster. This is the largest cluster among all the clusters in the network. Let $P_\lambda(B_\rho)$ be the probability of the event B_ρ under sensor density λ. We are interested in how $P_\lambda(B_\rho)$ changes as λ changes. To analyze this probability we first need to discuss some general properties of clusters of hexagons.

3.2 The largest cluster in the hexagon network

Here we first focus on a more coarse problem. For each hexagon, instead of counting how many sensors in it, we only ask if it is occupied. We then define an event similar to B_ρ as follows. Suppose C is a cluster of connected hexagons in the network. Let H_C be the total number of hexagons in cluster C. We define the relative size of C in terms of the relative area in the network:

$$r_H(C) = \frac{H_C}{M^2},$$

and define, for a given ρ, $1/2 < \rho < 1$, the event

$$D_\rho = \{\text{There is a cluster } C \text{ such that } r_H(C) \geq \rho\}. \quad (4)$$

Let $P_\lambda(D_\rho)$ be the probability of the event D_ρ under sensor density λ. We can also view $P_\lambda(D_\rho)$ as the probability that certain percentage of the whole sensor area is covered by the largest cluster of connected hexagons which are occupied by at least one sensor.

To analyze this probability, we introduce a set of binary random variables $X = \{X_{i,j} : \ i, \ j = 1, \cdots, M\}$ such that $X_{i,j} = 1$, if $H_{i,j}$ is occupied and $X_{i,j} = 0$ otherwise. Then $X_{i,j}$, i, $j = 1, \cdots, M$, are independent random variables with $P(X_{i,j} = 1) = 1 - P(X_{i,j} = 0) = p(\lambda)$. Let $\Omega = \{0,1\}^{\{1,\cdots,M\}\times\{1,\cdots,M\}}$. A realization $x = \{x_{i,j} : \ i, j = 1, \cdots, M\} \in \Omega$ of X defines a network configuration. Ω is the configuration space of the network.The event D_ρ is completely determined by the realizations $x \in \Omega$ of X.

If we use the center of each hexagon to represent that hexagon, and declare that a center is "open" if the corresponding hexagon is occupied, and "close" if otherwise, then it is not difficult to see that our problem here is actually a site percolation problem in a triangular lattice. A well known result from the percolation theory [Kesten 1982] asserts that there is a critical probability $p_0 = 1/2$, such that to have an infinite cluster in the lattice as $M \to \infty$, it is necessary and sufficient to have $p(\lambda) > p_0 = 1/2$. Since $p(\log 2/A_H) = 1/2$, an immediate consequence of this observation is the following proposition.

PROPOSITION 3.2.1. If $\lambda < \log 2/A_H$ then $\lim_M P_\lambda(D_\rho) = 0$.

PROOF. Let D_0 be the event that there is an infinite cluster of occupied hexagons in the infinite hexagon lattice. Then $P_\lambda(D_0) = 0$ when $\lambda < \log 2/A_H$. For every integer $M > 0$, let D_ρ^M be the event that D_ρ occurs in a finite box of size $M \times M$ of hexagons in the infinite hexagon lattice. Then $\{D_\rho^M \ i.o.\} \subset D_0$ and therefore $\limsup_{M\to\infty} P_\lambda(D_\rho) \leq P_\lambda(D_\rho^M \ i.o.) \leq P_\lambda(D_0) = 0$. The proposition follows. □

Some other simple and useful properties of $P_\lambda(D_\rho)$ are given in the following proposition.

PROPOSITION 3.2.2. For every $M < \infty$, $P_\lambda(D_\rho)$ is a differentiable and strictly increasing function of λ such that $P_0(D_\rho) = 0$ and $P_\infty(D_\rho) = 1$.

PROOF. Since $P_\lambda(D_\rho)$ can be written as a polynomial of $p(\lambda)$, the smoothness of this function in λ is trivial. To see that it is also increasing in λ, we note that the event D_ρ is an increasing event in the sense that if D_ρ occurs under some configuration $x' = \{x'_{i,j}\}$, then D_ρ will also occur under any configuration $x'' = \{x''_{i,j}\}$ satisfying $x''_{i,j} \geq x'_{i,j}$ for all i and j. A simple coupling argument in [Grimmett 1999] shows that the probability of the increasing event D_ρ is a strictly increasing function of $p(\lambda)$ which, in turn, is a strictly increasing function of λ. □

The following preliminary result is the key to the main theorems of this paper.

THEOREM 3.2.3. There is a constant $c > 0$, independent of M, and a $\lambda_0 = \lambda(A_H, M, \rho) > \log 2/A_H$ such that for all positive $\lambda \leq \lambda_0$,

$$P_\lambda(D_\rho) \leq \frac{1}{2} M^{-c[p(\lambda_0)-p(\lambda)]} \quad (5)$$

and, for all $\lambda \geq \lambda_0$, any small $\delta > 0$, and any small $\varepsilon_1 > 0$, there is an $M_0(\delta, \varepsilon_1)$ such that for all $M > M_0(\delta, \varepsilon)$,

$$P_\lambda(D_{\rho-\delta}) \geq 1 - \left(\frac{1}{2} + \varepsilon_1\right) M^{-c[p(\lambda)-p(\lambda_0)]}. \tag{6}$$

PROOF. See the proof in section IV. □

We call λ_0 in Theorem 3.2.3 the critical density. Note that λ_0 depends on M, ρ and, in particularly, A_H.

THEOREM 3.2.4. For a given ρ, the critical density $\lambda_0 = \lambda_0(\rho, M)$ has the property:

$$P_{\lambda_0}(D_\rho) \leq \frac{1}{2} \leq \liminf_{M\to\infty} P_{\lambda_0}(D_{\rho-\delta}), \tag{7}$$

for any small $\delta > 0$.

PROOF. It follows immediately from Theorem 3.2.3. □

Estimates similar to (5) and (6) can be obtained for the probability $P_\lambda(B_\lambda)$. They are given as the main theorems of this paper in the next subsection.

3.3 Probability of connectivity

For the connectivity probability $P_\lambda(B_\rho)$ and the critical density λ_0 we have:

THEOREM 3.3.1. Let λ_0 and constant c be as given in Theorem 3.2.3. For every fixed small δ and small ε_1, and for all sufficiently large M (depending on δ and ε_1),

$$P_\lambda(B_{\rho+\delta}) \leq \left(\frac{1}{2} + \varepsilon_1\right) M^{-c[p(\lambda_0)-p(\lambda)]}, \tag{8}$$

whenever $\lambda \leq \lambda_0$, and

$$P_\lambda(B_{\rho-\delta}) \geq 1 - \left(\frac{1}{2} + \varepsilon_1\right) M^{-c[p(\lambda)-p(\lambda_0)]}, \tag{9}$$

whenever $\lambda \geq \lambda_0$.

PROOF. See the proof in section IV. □

Intuitively, since for small δ, $P_\lambda(B_{\rho-\delta}) \approx P_\lambda(B_{\rho+\delta}) \approx P_\lambda(B_\rho)$, Theorem 3.3.1 asserts basically that, for sufficiently large M and some small $\varepsilon > 0$, if λ' and λ'' are such that $\lambda' < \lambda_0 < \lambda''$ and that $P_{\lambda'}(B_\rho) = \varepsilon$ and $P_{\lambda''}(B_\rho) = 1 - \varepsilon$, then the distance between λ'' and λ' is

$$\lambda'' - \lambda' = O\left(-\frac{\log \varepsilon}{\log M}\right).$$

In other words, if M is sufficiently large, a relatively small increment in sensor density λ in a neighborhood of λ_0 can result a significant increase of the probability $P_\lambda(B_\rho)$. On the other hand, any change of values of λ outside this neighborhood will have much less significant influence. This is called a *sharp-threshold property.* While the existence of λ_0 is established, its exact value is unknown. It should be determined experimentally.

The following assertion provides some information about where the critical sensor density is located.

THEOREM 3.3.2. For a given ρ, the critical density $\lambda_0 = \lambda_0(\rho, M)$ has the property:

$$\limsup_{M\to\infty} P_{\lambda_0}(B_{\rho+\delta}) \leq \frac{1}{2} \leq \liminf_{M\to\infty} P_{\lambda_0}(B_{\rho-\delta}), \tag{10}$$

for any small $\delta > 0$.

PROOF. This is a direct consequence of (8) and(9). □

A problem related to the connectivity problem is the coverage of sensor network. Since we do not require 100% connectivity of sensors in our study, it is necessary to consider only the coverage of sensors in an connected cluster. Let B_ρ^S be the event that at least $\rho * 100\%$ area is covered by the sensors from the largest cluster. The area covered by the sensors in a cluster depends on sensing radius R_s. As we pointed out before, in order to let two sensor nodes in neighboring hexagons be able to communicate with each other, the edge of hexagons should be no greater than $R_c/\sqrt{13}$. In order to let a sensor inside a hexagon be able to cover the entire hexagon area, R_s should be:

$$R_s \geq 2R_c/\sqrt{13}. \tag{11}$$

When the above inequality holds, we have $B_\rho \subset B_\rho^S$. In other words, good coverage can always be achieved with high probability as long as the same percentage of connectivity is achieved and (11) holds. We summarize this into the following theorem.

THEOREM 3.3.3. *$P_\lambda(B_\rho^S) \geq P_\lambda(B_\rho)$, whenever $R_s \geq 2R_c/\sqrt{13}$.*

Therefore (9) can also be used to estimate the lower bound of the coverage probability. Due to Theorem 3.2.3, we have a good reason to expect that the sensor coverage probability would have the similar sharp-threshold property. That is, for a given percentage of area coverage ρ, $P_\lambda(B_\rho^S)$ will have a sharp increase around a critical density. This property is verified by the simulation results.

3.4 Connectivity in the disk model

Now we discuss how to extend above results to the disk model which should be rather straightforward. The difference between the disk model and the hexagon model is the communication assumption between two nodes. In the hexagon model we assume two sensors are connected only if they are in the same hexagon or in neighboring hexagons. The distribution of the locations of the sensors in both cases are still the same. They both follow the same distribution of the Poisson point process. We can approximate the connectivity problem of the disk model with two hexagon models, H_1 and H_2 say, to obtain upper and lower bounds. This can be done as follows.

Let B_ρ^D be the event that at least a fraction $\rho\%$ of the sensors are connected in a single cluster in the disk model. Suppose the transmission range for all sensors is R_c. In the first hexagon model H_1, we scale the size of the hexagons so that the farthest distance between points of two neighboring hexagons equals R_c (see Figure 2(a)). Then, a connection between any two hexagons in the lattice always implies a connection between any pair of sensors inside these two hexagons in the disk model. Let $B_\rho^{H_1}$ be the event in H_1 as defined before. Then, $B_\rho^{H_1} \subset B_\rho^D$ and $P(B_\rho^{H_1}) \leq P(B_\rho^D)$. Similarly, we can scale the size of the hexagons so that whenever the distance between two sensors is less or equal to R_c, these two sensors are either in the same hexagon or in two neighboring hexagons (see Figure 2(b)). We choose the smallest possible hexagon size for this to happen and define the corresponding lattice as H_2. Let $B_\rho^{H_2}$ be the corresponding connectivity event. Then $P(B_\rho^D) \leq P(B_\rho^{H_2})$.

Applying Theorem 3.3.1 to both $B_\rho^{H_1}$ and $B_\rho^{H_2}$, we conclude that there are $\lambda_1^{H_1}$ and $\lambda_2^{H_2}$ such that (8) holds for $B_\rho^{H_1}$ and $\lambda_1^{H_1}$, and (9) holds for $B_\rho^{H_2}$ and $\lambda_2^{H_2}$.

THEOREM 3.4.1. There is a constant $c > 0$, independent of M, and two sensor densities λ_1^H and λ_2^H such that for every fixed small $\delta > 0$ and small ε_1, and for all sufficiently M,

$$P_\lambda(B_{\rho+\delta}) \leq \left(\frac{1}{2} + \varepsilon_1\right) M^{-c[p(\lambda_1^{H_1}) - p(\lambda)]}, \tag{12}$$

whenever $\lambda \leq \lambda_1^{H_1}$, and

$$P_\lambda(B_{\rho-\delta}) \geq 1 - \left(\frac{1}{2} + \varepsilon_1\right) M^{-c[p(\lambda) - p(\lambda_2^{H_2})]}, \tag{13}$$

whenever $\lambda \geq \lambda_2^{H_2}$.

We do not have a theoretical upper bound for the distance between $\lambda_1^{H_1}$ and $\lambda_2^{H_2}$ at this point, due to the techniques used in this paper. But simulation studies suggest strongly that (12) and (13) hold with $\lambda_1^{H_1}$ and $\lambda_2^{H_2}$ being replaced by a $\lambda_3 \in (\lambda_1^{H_1}, \lambda_2^{H_2})$. In other words, the critical sensor density should exist for the disk model as well.

4. PROOFS OF THE MAIN THEOREMS

In this section we prove Theorem 3.2.3 and Theorem 3.3.1. For the sake of easy reading, we restate the theorems before their proofs.

4.1 Proof of Theorem 3.2.3

Theorem 3.2.3. *There is a constant $c > 0$, independent of M, and a $\lambda_0 = \lambda(A_H, M, \rho) > \log 2/A_H$ such that for all positive $\lambda \leq \lambda_0$,*

$$P_\lambda(D_\rho) \leq \frac{1}{2} M^{-c[p(\lambda_0) - p(\lambda)]} \tag{14}$$

and, for all $\lambda \geq \lambda_0$, any small $\delta > 0$, and any small $\varepsilon_1 > 0$, there is an $M_0(\delta, \varepsilon_1)$ such that for all $M > M_0(\delta, \varepsilon)$,

$$P_\lambda(D_{\rho-\delta}) \geq 1 - \left(\frac{1}{2} + \varepsilon_1\right) M^{-c[p(\lambda) - p(\lambda_0)]}. \tag{15}$$

A key part of the proof is a result called sharp-threshold inequality ([Bourgain et al. 1992], [Friedgut and Kalai 1996], and [Graham and Grimmett 2006]) for the product probability measures. To apply the result we proceed as follows.

We first make the hexagonal lattice $H = \{H_{i,j},\ i,j = 1, \cdots, M\}$ into a torus $H^* = \{H_{i',j'},\ i',j' \in \mathbf{Z}\}$ by identifying $H_{i',j'} \in H^*$ with $H_{i,j} \in H$ whenever $i' = i \bmod M$ and $j' = j \bmod\ M$. In this torus we define for every k and $l \in \mathbf{Z}$ a shift translations $\tau_{k,l} :\ H_{i,j} \rightarrow H_{i+k,j+l},\ i,j \in \mathbf{Z}$. Then $\tau = \{\tau_{k,l},\ k,l \in \mathbf{Z}\}$ forms a subgroup of automorphisms on the torus with the transitivity property: any hexagon $H_{i,j}$ can be shifted to any other hexagon $H_{i',j'}$ with the translation $\tau_{i'-i,j'-j}$.

Next, we extend the definition of clusters of the hexagons in the torus in a natural way. This means that the hexagons in the 1st row can be connected to the hexagons in the Mth row, and the 1st column of the hexagons can be connected to the Mth column of hexagons, depending on weather the corresponding hexagons are occupied or not. We then extend the definition of connectivity of any two hexagons into the torus accordingly. Of course, two hexagons joined in the torus need not be connected in the original lattice H.

Let D^*_ρ be the event that at least a fraction $\rho\%$ of the occupied hexagons are connected *in the torus*. Clearly, $D_\rho \subset D^*_\rho$ but $D_\rho \neq D^*_\rho$, that is, the occurrence of D^*_ρ does not necessary imply the occurrence of D_ρ. This is because in the torus, the connections between hexagons can go through the boundary edges of the original lattice to reach the hexagons on the other side. If such connections are the only ones making D^*_ρ to occur, D_ρ will not occur. The reason we introduce D^*_ρ is that it has some nice properties which are necessary for our proof. The connection between D^*_ρ and D_ρ will become clear later.

It is not difficult to see that D^*_ρ satisfies the following properties: (a) it is an increasing event like D_ρ (see the proof of Proposition 2) and (b) it is invariant under the shift translations $\tau_{k,l}$. The later property means if D^*_ρ occurs under a network configuration $x' \in \Omega$, and if $x'' \in \Omega$ is any shift translation of x', then D^*_ρ also occurs under the configuration x''. We note that property (b) of D^*_ρ can not be defined for D_ρ.

Let $P_\lambda(D^*_\rho)$ be the probability that D^*_ρ will occur. In virtue of properties (a) and (b) of D^*_ρ, the sharp-threshold inequality in [Graham and Grimmett 2006] implies:

Lemma 4.1.1. There is a constant $c > 0$, independent of M and λ, such that

$$\frac{d}{d\lambda} P_\lambda(D^*_\rho) \geq c^*(\lambda) \min\{P_\lambda(D^*_\rho), 1 - P_\lambda(D^*_\rho)\} \log M \qquad (16)$$

with $c^*(\lambda) = cA_H e^{-\lambda A_H}$.

Proof. Since this is a consequence of a more general result, we only outline the main steps of the proof to show the connections among the key concepts. For more details we refer the readers to [Graham and Grimmett 2006] and the references therein.

Define the “conditional influence” of each hexagon $H_{i,j}$ on the event D^*_ρ as

$$I_{D^*_\rho}(i,j) = P_\lambda(D^*_\rho | X_{i,j} = 1) - P_\lambda(D^*_\rho | X_{i,j} = 0).$$

Then there is a constant $c > 0$, independent of M and λ, and a hexagon $H_{i,j}$, such that

$$I_{D^*_\rho}(i,j) \geq c \min\{P_\lambda(D^*_\rho), 1 - P_\lambda(D^*_\rho)\} \log M/M^2.$$

Since both the event D^*_ρ and the product measure $P_\lambda(\cdot)$ are invariant under the translation group τ, it is necessary that

$$I_{D^*_\rho}(i,j) = I_{D^*_\rho}(i',j') \text{ for all } i,j,i',j' \in \mathbf{Z}.$$

On the other hand, let $p = p(\lambda) = 1 - e^{-\lambda A_H}$ and treat the quantity $P_\lambda(D^*_\rho)$ as a function of p, then we have, from formula (1.4) of [Graham and Grimmett 2006], the identity

$$\frac{d}{dp} P_\lambda(D^*_\rho) = \sum_{i,j} I_{D^*_\rho}(i,j).$$

Therefore

$$\frac{d}{d\lambda} P_\lambda(D^*_\rho) = \frac{dp}{d\lambda} \times \frac{d}{dp} P_\lambda(D^*_\rho) = A_H e^{-\lambda A_H} \sum_{i,j} I_{D^*_\rho}(i,j).$$

The Lemma follows from these observations. □

As a consequence of this lemma, we have

LEMMA 4.1.2. There is a $\lambda_0 > 0$ depending on M and ρ such that:

$$P_\lambda(D^*_\rho) \le \frac{1}{2} M^{-c(p(\lambda_0)-p(\lambda))}, \quad \text{for } 0 \le \lambda \le \lambda_0 \tag{17}$$

and

$$P_\lambda(D^*_\rho) \ge 1 - \frac{1}{2} M^{-c(p(\lambda)-p(\lambda_0))}, \quad \text{for } \lambda_0 \le \lambda < \infty, \tag{18}$$

where c is the constant given in Lemma 4.1.1.

PROOF. We can argue as in the proof of Proposition 3.2.2 that $P_\lambda(D^*_\rho)$ is continuous and increasing in λ with minimum 0 and maximum 1, and therefore there is a λ_0, $0 < \lambda_0 < \infty$, such that

$$P_{\lambda_0}(D^*_\rho) = 1/2. \tag{19}$$

It also follows from the increasing property of $\lambda \to P_\lambda(D^*_\rho)$ that for every $\lambda \le \lambda_0$, $P_\lambda(D^*_\rho) \le 1 - P_\lambda(D^*_\rho)$, and for every $\lambda \ge \lambda_0$,$P_\lambda(D^*_\rho) \ge 1 - P_\lambda(D^*_\rho)$. Therefore (16) takes the form:

$$\frac{d}{d\lambda} P_\lambda(D^*_\rho) \ge c^*(\lambda) P_\lambda(D^*_\rho) \log M, \quad \text{if } \lambda \le \lambda_0$$

and

$$\frac{d}{d\lambda} P_\lambda(D^*_\rho) \ge c^*(\lambda)[1 - P_\lambda(D^*_\rho)] \log M, \quad \text{if } \lambda \ge \lambda_0.$$

The inequalities can be further written as

$$\frac{d}{d\lambda} \log P_\lambda(D^*_\rho) \ge c^*(\lambda) \log M, \quad \text{if } \lambda \le \lambda_0$$

and

$$-\frac{d}{d\lambda} \log(1 - P_\lambda(D^*_\rho)) \ge c^*(\lambda) \log M \quad \text{if } \lambda \ge \lambda_0.$$

Integrating both sides of these two inequalities from λ to λ_0 and from λ_0 to λ respectively, and applying (19), we obtain (17) and (18) □

Let λ_0 in Theorem 3.2.3 be as given in Lemma 4.1.2. To show (14) of theorem 3.2.3 we note that since event D_ρ implies event D^*_ρ,

$$P_\lambda(D_\rho) \le P_\lambda(D^*_\rho). \tag{20}$$

The estimate (14) now follows from (17) of Lemma 4.1.2.

We next show (15). As it was pointed out above, there are cases where the occurrence of D^*_ρ does not imply the occurrence of D_ρ. To exclude such possibilities we consider a slightly large event $D_{\rho-\delta}$ for a small $\delta > 0$, and restrict D^*_ρ to some subset of the configurations in the following so that the occurrence of D^*_ρ in this subset will imply the occurrence of $D_{\rho-\delta}$ with a large probability.

Let $\phi(M)$ be any integer depending on M such that $\phi(M) \to \infty$, as $M \to \infty$, and

$$\phi(M) = o\left(\min\{c(p(\lambda) - p(\lambda_0)), 1\} \log M\right). \tag{21}$$

Let H_i, $i = 1, 2, 3, 4$, be the sub-lattices of H which are, respectively, the top, bottom, left and right boundary strips of the sizes $\phi(M) \times M$, $\phi(M) \times M$, $M \times \phi(M)$ and $M \times \phi(M)$. More specifically, we let

$$H_1 = \{H_{i,j} : i = M - \phi(M) + 1, \cdots M, j = 1, \cdots M\},$$

$$H_2 = \{H_{i,j} : i = 1, \cdots \phi(M), j = 1, \cdots M\},$$

$$H_3 = \{H_{i,j} : i = 1, \cdots , M, j = 1, \cdots , \phi(M)\},$$

$$H_4 = \{H_{i,j} : i = 1, \cdots M, j = M - \phi(M) + 1, \cdots M\}.$$

We are going to show that for each strip, if $\lambda > \lambda_0$ and if M is sufficiently large, then with a large probability there will be a path of occupied hexagons which connects one shorter side of the strip to the other shorter side of the strip (traversed long-way) within that strip. More precisely, if we let E_i be the event that there is a connected path of occupied hexagons inside H_i which crosses the strip H_i long-way, then we have

LEMMA 4.1.3. For $i = 1, 2, 3, 4$, there are constants $c_i > 0$ such that for all large M and $\lambda \ge \lambda_0$,

$$P_\lambda(E_i) \ge 1 - e^{-c_i \phi(M)}. \tag{22}$$

PROOF. The duality property of hexagon lattice and its consequences in the percolation problem is important in our proof, and this is why we specifically choose hexagon lattice for our model. We only need to point out the following key facts.

First, according to the duality property of hexagon lattice, the existence of a long-way crossing of a path of *occupied* hexagons in a rectangular region is equivalent to the non-existence of a short-way crossing of a path of *non-occupied* hexagons.

Second, for sufficiently large M, it is necessary that $p(\lambda_0)$ is larger than $p_c = 1/2$, the critical probability for the existence of an infinite cluster of occupied hexagons in the lattice. Therefore the probability for a hexagon to be non-occupied must be less p_c, which is also the critical probability for the existence of an infinite cluster of non-occupied hexagons.

The Lemma now can be proved with the standard arguments from the percolation theory in [Grimmett 1999] (the proof of Theorem 6.1 in page 118 and the comments between Lemma 11.21 and 11.22 in page 294). □

We are ready to estimate $P_\lambda(D_{\rho-\delta})$ for $\lambda > \lambda_0$ and any given small $\delta > 0$. Let $E = E_1 \cap E_2 \cap E_3 \cap E_4$, and consider the event $F = D^*_\rho \cap E$. Since $P_\lambda(F) = P_\lambda(F \cap D_{\rho-\delta}) + P_\lambda(F - D_{\rho-\delta})$, we have

$$P_\lambda(D_{\rho-\delta}) \geq P_\lambda(F) - P_\lambda(F - D_{\rho-\delta}). \tag{23}$$

Since D^*_ρ and E are both increasing events, a standard application of the Fortuin - Kasteleyn - Ginibre inequality in [Grimmett 1999] yields:

$$P_\lambda(F) \geq P_\lambda(D^*_\rho)P^2_\lambda(E_1)P^2_\lambda(E_3) \tag{24}$$

(since $P_\lambda(E_1) = P_\lambda(E_2)$ and $P_\lambda(E_3) = P_\lambda(E_4)$). Therefore it follows from Lemma 4.1.3 that there is a $b > 0$ such that for all sufficiently large M,

$$P_\lambda(F) \geq 1 - \frac{1}{2}M^{-c(p(\lambda)-p(\lambda_0))} - O(e^{-b\phi(M)}). \tag{25}$$

Because of (21), this implies that for any given $\varepsilon_1 > 0$ and sufficiently large M, depending on ε_1,

$$P_\lambda(F) \geq 1 - \left(\frac{1}{2} + \varepsilon_1\right) M^{-c(p(\lambda)-p(\lambda_0))}. \tag{26}$$

We now claim that $F - D_{\rho-\delta} = \emptyset$, and therefore $P(F - D_{\rho-\delta}) = 0$ in (23), for all large M. Indeed, the occurrence of F implies that there is a close circuit of the path of occupied hexagons surrounding the hexagons in the sub-lattice $H - H_1 \cup H_2 \cup H_3 \cup H_4$. It follows that there must be a cluster connected *within* the original lattice (without the torus structure) containing at least $\rho M^2 - 4\phi(M)(M - \phi(M))$ hexagons, where $4\phi(M)(M - \phi(M))$ is the total number of hexagons in the strips H_i, $i = 1, 2, 3, 4$. In other words, $F \subset D_{\rho-\delta_1}$ with $\delta_1 = 4\phi(M)(M - \phi(M))/M^2$. Therefore when M is so large that $\delta_1 < \delta$, $F \subset D_{\rho-\delta}$. We obtain (15) from (23) and (26).

This completes the proof of Theorem 3.2.3.

4.2 Proof of Theorem 3.3.1

Theorem 3.3.1. *Let λ_0 and constant c be as given in Theorem 3.2.3. For every fixed small δ and small ε_1, and for all sufficiently large M (depending on δ and ε_1),*

$$P_\lambda(B_{\rho+\delta}) \leq \left(\frac{1}{2} + \varepsilon_1\right) M^{-c[p(\lambda_0)-p(\lambda)]}, \tag{27}$$

whenever $\lambda \leq \lambda_0$, and

$$P_\lambda(B_{\rho-\delta}) \geq 1 - \left(\frac{1}{2} + \varepsilon_1\right) M^{-c[p(\lambda)-p(\lambda_0)]}, \tag{28}$$

whenever $\lambda \geq \lambda_0$.

The proof is based on a simple application of a large deviation argument. Let us first consider the case when $\lambda \leq \lambda_0$. Since $P_\lambda(B_{\rho+\delta}) = P_\lambda(B_{\rho+\delta} \cap D_\rho) + P_\lambda(B_{\rho+\delta} - D_\rho)$, we have

$$P_\lambda(B_{\rho+\delta}) \leq P_\lambda(D_\rho) + P_\lambda(B_{\rho+\delta} - D_\rho).$$

To show (27) we only need to show, in virtue of Theorem 3.2.3, that $P_\lambda(B_{\rho+\delta} - D_\rho)$ is of the order $o\left(M^{-c[p(\lambda_0)-p(\lambda)]}\right)$. In fact, it is not difficult to show that when δ is fixed, this quantity decays to 0 at a much faster rate.

For every configuration $x \in \Omega$, let $\mathcal{C}(x) = \{C_1, \cdots, C_K\}$ be the set of clusters in x. Suppose $\mathcal{C}(x)$ and $N_A \sim Poisson(\lambda A)$ are given. Then, for $i = 1, \cdots, K$, the conditional distribution of N_{C_i}, the number of sensors in C_i, given $\mathcal{C}(x)$ and N_A, the number of sensors in the field, is the binomial distribution:

$$N_{C_i} \mid \mathcal{C}(x), N_A \sim B(r_A(C_i), N_A).$$

Let C_{i_0} be a cluster in $\mathcal{C}(x)$ such that $r_N(C_{i_0}) = N_{C_{i_0}}/N_A \geq \rho + \delta$. Since the event D_ρ^c implies that $r_A(C_i) < \rho$ for all i, we have

$$B_{\rho+\delta} - D_\rho \subset \{r_N(C_{i_0}) \geq \rho + \delta, r_A(C_{i_0}) < \rho\}.$$

The standard large deviation result in [Durrett 1991] implies that there is an $\alpha(\rho, \delta) > 0$ such that

$$P(r_N(C_{i_0}) \geq \rho + \delta | r_A(C_{i_0}) < \rho, N_A) \leq e^{-\alpha(\rho,\delta)N_A}.$$

It follows that

$$\begin{aligned}
&P_\lambda(B_{\rho+\delta} - D_\rho) \\
&\leq P(r_N(C_{i_0}) \geq \rho + \delta, r_A(C_{i_0}) < \rho) \\
&= E[P(r_N(C_{i_0}) \geq \rho + \delta | \mathcal{C}(x), N_A),\ r_A(C_{i_0}) < \rho] \\
&\leq E\left(e^{-\alpha(\rho,\delta)N_A}\right) \\
&= \exp\left(-\lambda A(1 - e^{-\alpha(\rho,\delta)})\right).
\end{aligned}$$

Recall that $A = O(M^2)$. Therefore for every fixed $\delta > 0$, $P_\lambda(B_{\rho+\delta} - D_\rho)$ decays to 0 at a far more faster rate than M^{-d} for any finite $d > 0$. This shows (27).

The case of $\lambda \geq \lambda_0$ can be proved with basically the same arguments and therefore (28) holds.

5. SIMULATIONS

The results above are obtained mainly for the hexagon model where nodes are assumed to have communication links with other nodes in the same hexagon or neighboring hexagons. A more commonly used communication model for sensor networks is the disk model where two nodes within each other's transmission range have a communication link. The purpose of the current simulations is two-fold. First, we study the connectivity probability in both the hexagon model and the disk model, in order to verify the theoretical results for the disk model. Second, we study how the connectivity probability, under various connectivity percentage ρ, is affected by the sensor density λ and the average number of neighbor nodes N_b.

5.1 Methodology and Simulation Settings

The simulation code is written in *Matlab*. The sensor field is a square region with an area of 50×50 units. We would like to remark that this is actually an appropriate size for the area. We do not use any larger sensor field in our simulations, since this

```
Input: ρ.
Output: Sequence of pairs (λ,P_λ(B_ρ)).
λ = Min Value;
REPEAT
   cnt = 0;
  FOR i = 1 to n
      Generate a Poisson number N_A with mean λ;
      Place N_A nodes randomly in the area;
      Obtain a network by connecting nodes using hexagon model (or disk model);
      IF exist a cluster larger than 100ρ%
         cnt++;
      ENDIF
   ENDFOR
   P_λ(B_ρ)=cnt/n;
   Output (λ, P_λ(B_ρ));
   λ+ = h; // h is a small increment of λ
UNTIL λ > Max Value
```

Fig. 1. Algorithm

study is not about a limiting property of sensor connectivity in which the area of the sensor field goes to infinity, but rather, we are interested in sensor fields of any reasonable size (preferably, smaller sizes) at which the sharp-threshold phenomenon emerges. The sharp-threshold property can only get more and more sharper as the size of the sensor field gets larger and larger. The number N_A of sensors follows the Poisson distribution with given density λ. The N_A sensor nodes are randomly placed in the sensor field. All sensor nodes have a fixed transmission range R_c, which is set to 1 throughout the simulations. In disk model, we assume two nodes can communicate with each other if and only if the distance between them is smaller than R_c; The connectivity of sensors within the hexagon model is defined as before. The size of the hexagons will be scaled according to R_c (see the next subsection).

The algorithm, as shown in Figure 1, takes the connectivity percentage ρ as input and outputs a sequence of pairs of λ and $P_\lambda(B_\rho)$ (the probability of meeting ρ at density λ). To compute $P_\lambda(B_\rho)$, we generate a sufficiently large number n of network samples to obtain a stable estimate of $P_\lambda(B_\rho)$ which is the total number of network samples that meet the percentage ρ divided by n. The plots in all simulation figures are the average of 100 runs.

5.2 Hexagon Size in the Hexagon Model

In this set of simulations, we compare the connectivity probability of the disk model to those of several hexagon models.

In the first hexagon model, we set the size of hexagons to be as large as possible as long as any connected nodes in the hexagon model remains connected in the disk model. In this case, the maximum distance between two nodes in neighboring hexagons should be smaller than the transmission range R_c in disk model. Thus, the length of the hexagon edge l should be equal to $R_c/\sqrt{13}$. This gives the lower bound for the hexagon size. See Figure 2(a). In the second hexagon model, we set the size of hexagons to be as small as possible, as long as any connected nodes in the disk model remain connected in the hexagon model. It is easy to see that this

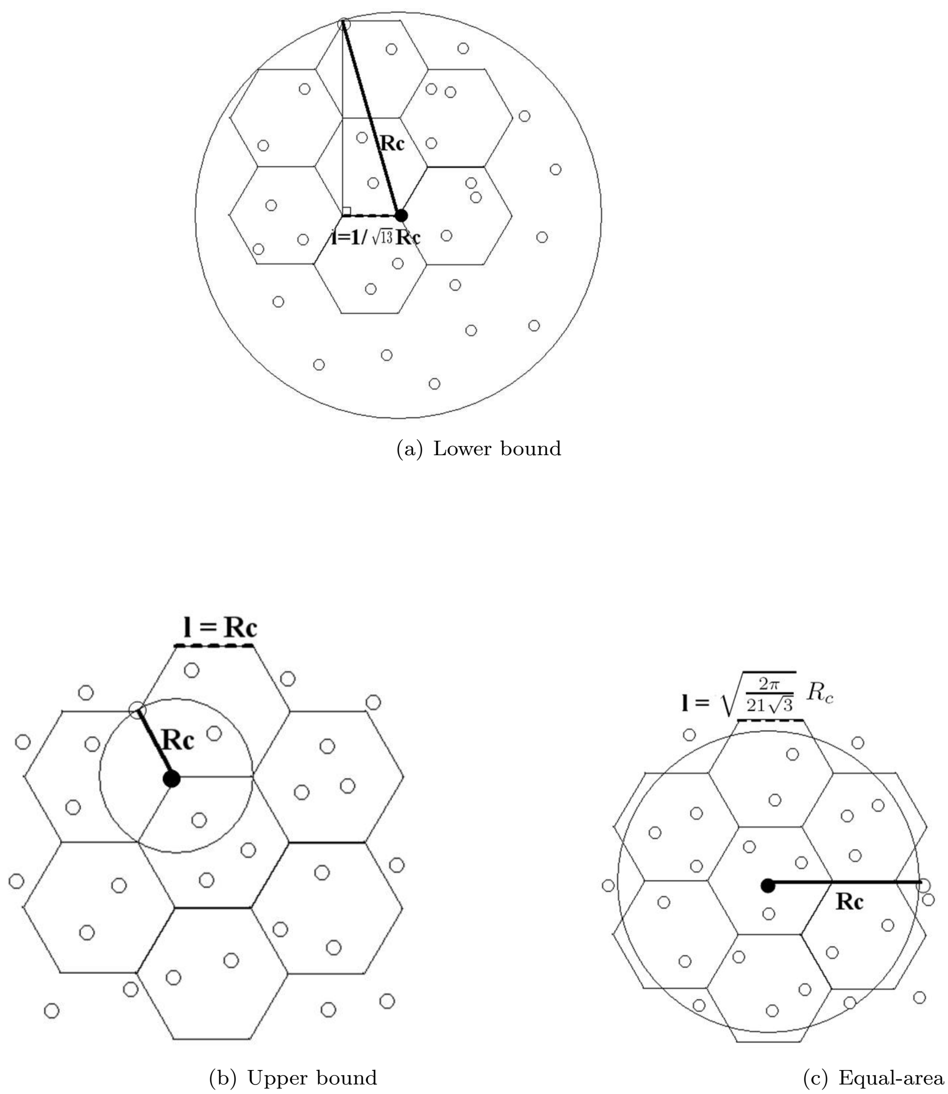


Fig. 2. Lower bound, upper bound and equal-area of hexagon sizes

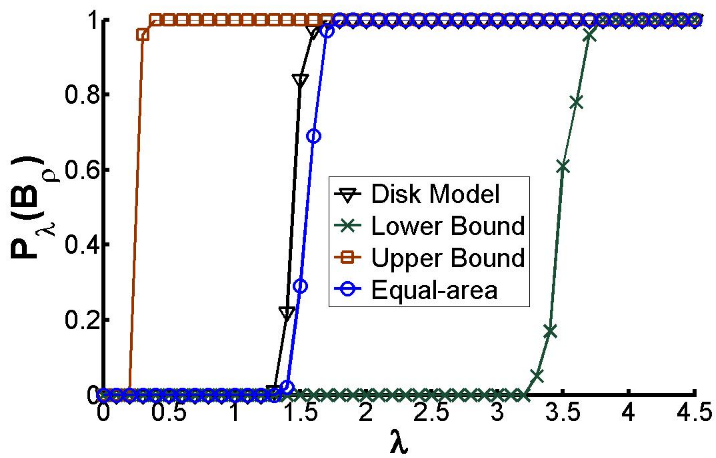


Fig. 3. $P_\lambda(B_\rho)$ versus λ with different hexagon sizes

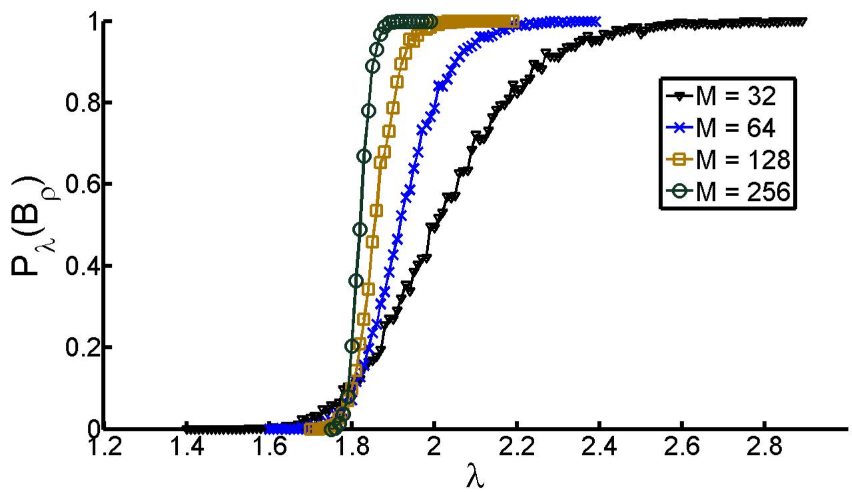
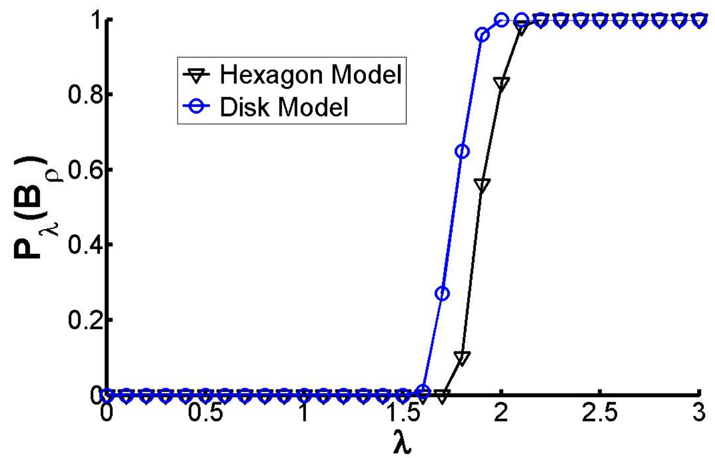


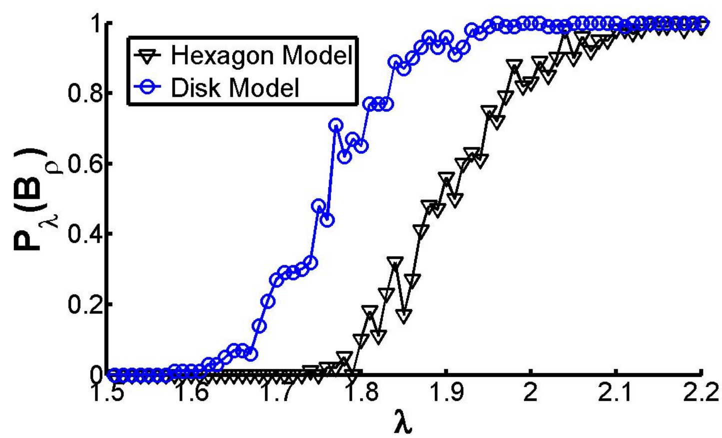
Fig. 4. $P_\lambda(B_\rho)$ versus λ with different M

requires the length of hexagons $l = R_c$. This gives the upper bound for the hexagon size. See Figure 2(b). We also study a third hexagon model, called an equal-area model, in which the size of the hexagon is such that the total area of any hexagon along with its 6 neighbors equals the area covered by a disk with radius R_c. Figure 2(c) illustrates the case. The length of hexagon edge l is equal to $\sqrt{\frac{2\pi}{21\sqrt{3}}}\, R_c$ in this case.

Figure 3 shows the probability $P_\lambda(B_\rho)$ versus λ for the disk model and three cases of the hexagon model. From Figure 3 we observe the following. 1) The hexagon model with various hexagon sizes exhibits the same pattern as the disk model. There is always a sharp increase of $P_\lambda(B_\rho)$ from 0 to 1 near some critical



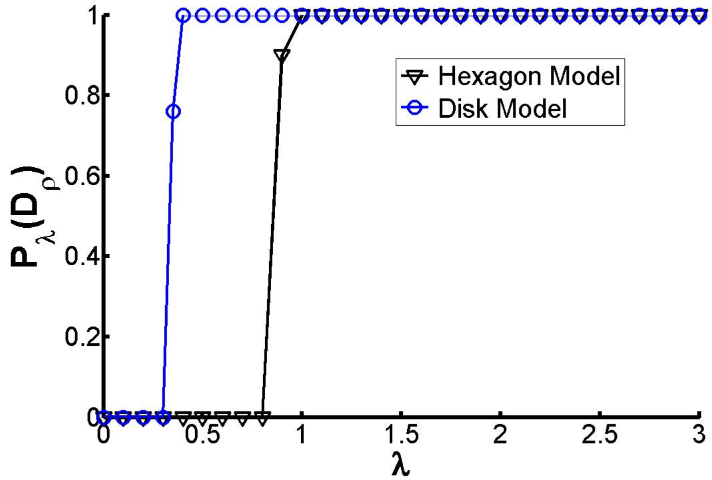
(a) $h = 0.1$



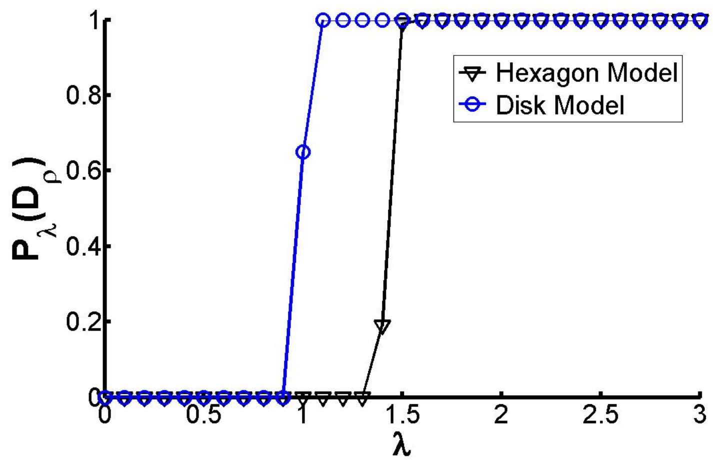
(b) $h = 0.01$

Fig. 5. $P_\lambda(B_\rho)$ versus λ for equal-area hexagon model when $\rho = 0.95$

density λ_0. 2) The values of λ_0 for different size of hexagons have a large variance. As the hexagon size increases, λ_0 decreases from 3.5 (for lower bound of hexagon size) to 0.2 (for upper bound of hexagon size). This is because for a larger hexagon size, there is a higher chance for nodes to fall into the same hexagon or neighboring hexagons resulting a higher probability of network connection. That is, to reach the same $P_\lambda(B_\rho)$, the hexagon model with a larger hexagon size requires less node density, leading to a smaller λ_0. 3) The critical density value of the equal-area hexagon model is very close to the disk model. For the rest of simulation, we use the hexagon size of equal-area for the hexagon model.



(a) $\rho = 0.65$



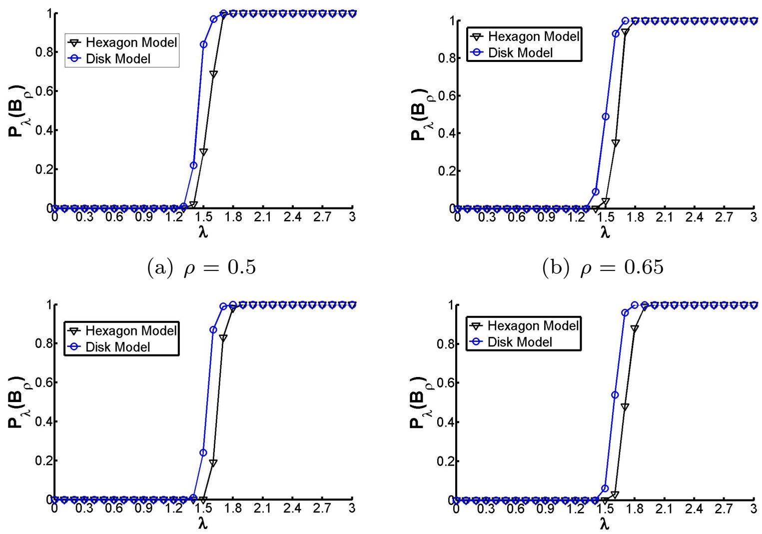
(b) $\rho = 0.95$

Fig. 6. $P_\lambda(D_\rho)$ versus λ

5.3 Sharp-Threshold Phenomenon Near Critical Density

Figure 4 shows the probability $P_\lambda(B_\rho)$ versus node density λ in the hexagon model with $\rho = .95$ and M =32, 64, 128, and 256. We notice that the threshold width of the probability from the simulation doubles as $\log M$ doubles. This is consistent with our theoretical finding which implies that the threshold width is of order $\log M$.

Figure 5 shows the $P_\lambda(B_\rho)$ versus node density λ in the hexagon and the disk models, where ρ is fixed at 0.95. In Figure 5(a), the increment of λ is set to be 0.1. To have a closer look at the interval of λ that has the sharp increase of $P_\lambda(B_\rho)$, we set the increment to 0.01 and obtain Figure 5(b). From Figure 5(b), we see that the two curves (for the hexagon and disk models) are similar with each other.



(a) $\rho = 0.5$ (b) $\rho = 0.65$

(c) $\rho = 0.75$ (d) $\rho = 0.85$

Fig. 7. $P_\lambda(B_\rho)$ versus λ with different ρ.

According to the previous analysis,λ_0 appears at a point where $P_\lambda(B_\rho)$ is nearly 1/2, and the interval of λ that causes the sharp increase of $P_\lambda(B_\rho)$ is $C/\log M$ for some constant. In our simulations, we place about 80 hexagons in each row to cover the 50×50 area, thus $1/\log 80$ is equal to 0.53. From Figure 5(b) we see that the interval for $P_\lambda(B_\rho)$ to increase from 0 to 1 is less than 0.53 for both the hexagon model and the disk model, suggesting $C < 1$. In terms of the number of sensor nodes deployed in the network, for the disk model, it has around 4127 sensors ($\lambda = 1.61$) when $P_\lambda(B_\rho)$ is still almost 0; but with 4914 sensors ($\lambda = 1.95$), $P_\lambda(B_\rho)$ is increased to almost 1. Similarly for the hexagon model, with the increase of number of sensors from 4409 ($\lambda = 1.75$) to 5408 ($\lambda = 2.14$), $P_\lambda(B_\rho)$ increases from 0 to 1. That is, with less than 20% of additional sensors, $P_\lambda(B_\rho)$ can be increased from 0 (no chance of connected) to 1 (highly sure of connected). From this result we see that it is important to find this critical density for large scale deployment of sensor networks. With a slight increase of node density over this critical point, the connectivity probability can be improved significantly.

In the theoretical analysis, we also proved the existence of critical density for sensor coverage and the sharp increase of probability for meeting the required percentage of sensor coverage. We use $P_\lambda(D_\rho)$ to denote the probability that the percentage of sensor area covered by connected sensors is greater than ρ under density λ. For the hexagon model, $P_\lambda(D_\rho)$ is simply calculated as the number of connected hexagons divided by the total number of hexagons, because all hexagons have the same size. The sensing range is assumed to be the outside diameter of hexagons. That is, a hexagon is fully covered whenever there is a sensor falling into the hexagon. For the disk model, we use Monte Carlo method with 100,000 random samples to compute the percentage of area covered by the connected sensors.

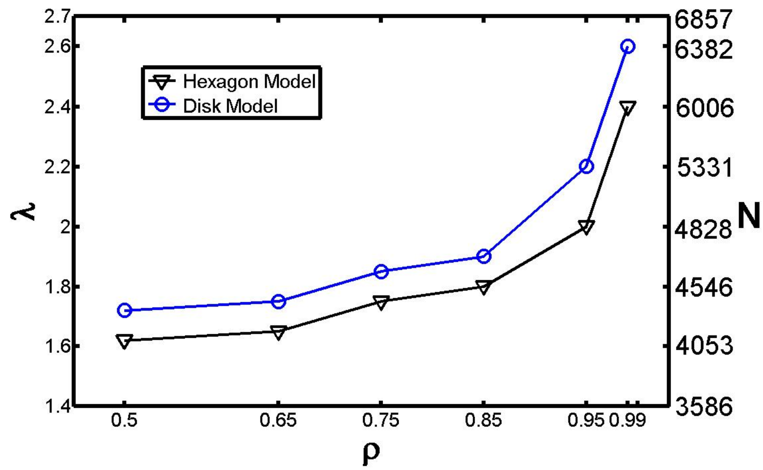


Fig. 8. ρ versus λ when $P_\lambda(B_\rho) \geq 0.99$

Figure 6(a) shows $P_\lambda(D_\rho)$ versus λ, where ρ is fixed at 0.65 (i.e., at least 65% of the area is covered by connected sensors). Figure 6(b) shows the case where ρ is fixed at 0.95. From Figure 6 we see that $P_\lambda(D_\rho)$ exhibits the same trend as $P_\lambda(B_\rho)$ at critical density points.

5.4 Impact of Parameter ρ

In this set of simulations, we investigate the impact of the parameter ρ. Figure 7 shows a group of charts of $P_\lambda(B_\rho)$ versus λ where ρ varies from 0.5 to 0.85 (Figure 5 shows the case where ρ is 0.95). All charts for different ρ exhibit the same pattern of the increase of $P_\lambda(B_\rho)$ as the increase of λ.

Figure 8 shows the correlation between ρ and λ when $P_\lambda(B_\rho) \geq 0.99$ (i.e., highly sure that $100\rho\%$ of nodes are connected). In general, we can see that the increase of λ with ρ is slow for small ρ. But, the increase of λ becomes sharper for large ρ. This is particularly true when ρ reaches 95% or even higher. Figure 8 also translates the density λ into the number of sensor nodes (the right hand side vertical bar). We can see when ρ is increased from 0.95 to 0.99, the number of sensors required to meet the connectivity percentages increases from 4828 to 6006, which is much higher than the increase of ρ. Therefore, in this case it is much more economic to sacrifice a small percentage of node connection to save a large number of sensor nodes.

5.5 The Number of Neighbors and Connectivity

As it was pointed out in related work, a number of works studied the relationship between the probability of full-connectivity and the average number of neighbors of nodes. In this set of simulations, we study the impact of the average number of connected neighbors N_b to the connectivity percentage of the network ρ and the connectivity probability $P_\lambda(B_\rho)$ in the hexagon model. Figure 9 shows $P_\lambda(B_\rho)$ versus N_b, where ρ is set to be 0.99 in the set of simulations. From Figure 9, we can see the same trend of the sharp increase of $P_\lambda(B_\rho)$ at some critical interval of

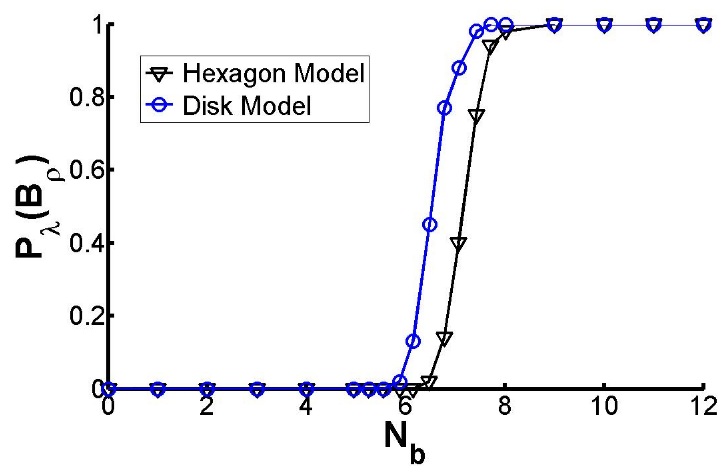


Fig. 9. $P_\lambda(B_\rho)$ versus N_b when $\rho = 0.99$

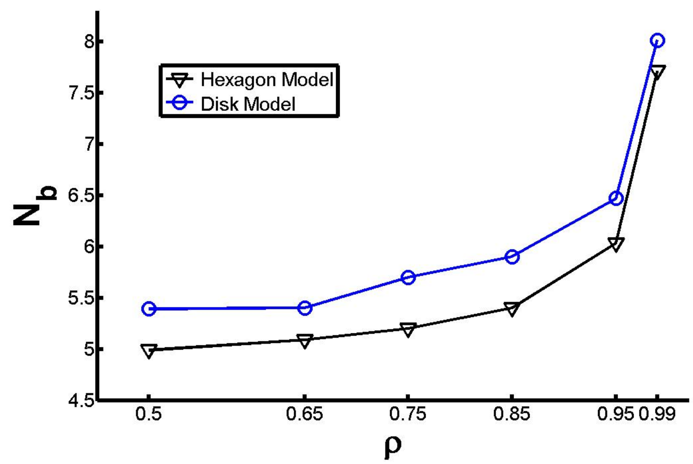


Fig. 10. ρ versus N_b when $P_\lambda(B_\rho) \geq 0.99$

N_b. $P_\lambda(B_\rho)$ increases sharply from 0 to 1 during a small interval of N_b between 5.8 and 7.4 for disk model and between 6.4 and 8.0 for hexagon model. This result is consistent with results obtained in [Ni and Chandler 1994] [Gilbert 1961] [Philips et al. 1989].

Figure 10 plots the correlation between ρ and N_b when $P_\lambda(B_\rho)$ is set to be greater than 0.99. The curves in this figure are surprisingly similar to those in Figure 8. This tells us the average number of neighbors of nodes is another metric that can describe the node density of a network accurately (which is equivalent to the metric λ). From Figure 10, we can also see that the increasing rate of N_b is faster when ρ is becoming larger, which confirms our conclusion that achieving a high ρ, it requires a much larger number of sensors. This is not economic for many sensor network applications.

6. CONCLUSION

In this paper, we present results on how the connectivity (or coverage) probability changes as the sensor density increases for the case of partial connectivity (or partial coverage). Through theoretical analysis and simulation study we made the following discoveries: 1) For a partial connectivity requirement ρ, $0.5 < \rho < 1$, there is a critical sensor density around which there is a sharp increase of the probability that at least a fraction $\rho\%$ of sensors are connected in the network. This sharp increase is from almost 0 to almost 1 within a short interval of sensor density. 2) The interval for this sharp increase is in the order of $O(1/\log A)$, where A is the sensor area, and the location of the critical density is the density at which the above probability is about 1/2. This result allows us to accurately determine the sensor density in order to meet the connectivity requirement with high probability. 3) We obtained the similar results of sharp increase of probability for partial sensor area coverage as the increase of sensor density.

REFERENCES

Bai, X., Kumar, S., Xuan, D., Yun, Z., and Lai, T. 2006. Deploying wireless sensors to achieve both coverage and connectivity. In *Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing(MobiHoc)*. ACM.

Bai, X., Zhang, C., Xuan, D., and Jia, W. 2009. Full-coverage and k-connectivity (k=14, 6) three dimensional networks. In *Proceedings of 28th IEEE Conference on Computer Communications (INFOCOM)*. IEEE.

Bettstetter, C. 2002. On the minimum node degree and connectivity of a wireless multihop network. In *Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing(MobiHoc)*. ACM.

Bourgain, J., Kahn, J., Kalai, G., Katznelson, Y., and Linial, N. 1992. The influence of variables in product spaces. *Israel Journal of Mathematics. 77,* 1-2, 55–64.

Cai, H., jia, X., and Sha, M. 2010. Critical sensor density for partial connectivity in large area wireless sensor networks. In *Proceedings of 29th IEEE Conference on Computer Communications (INFOCOM) mini-conference*. IEEE.

Dousse, O., Franceschetti, M., , and Thiran, P. 2006. A case for partial connectivity in large wireless multi-hop networks. In *Proceedings of Information Theory and Applications Workshop (UCSD-ITA)*.

Durrett, R. 1991. *The TEXProbability: Theory and Examples*. Duxbury.

Friedgut, E. and Kalai, G. 1996. Every monotone graph property has a sharp threshold. *Proceedings of the American Mathematical Society. 124,* 10, 2993–3002.

Gilbert, E. 1961. Random plane networks. *Journal of the Society for Industrial and Applied Mathematics. 9,* 4, 533–543.

Graham, B. and Grimmett, G. 2006. Influence and sharp-threshold theorems for monotonic measures. *The Annals of Probability. 34,* 5, 1726–1745.

Grimmett, G. 1999. *The TEXPercolation*. Springer-Verlag.

Hou, T. and Li, V. 1986. Transmission range control in multihop packet radio networks. *IEEE Trans. on Communications. 34,* 1, 38–44.

Kesten, H. 1982. *The TEXPercolation theory for mathematicians*. Birkhauser.

Kleinrock, L. and Silvester, J. 1978. Optimal transmission radii for packet radio networks or why six is a magic number. In *Proceedings of National Telecommunications Conference (NTC)*.

Meester, R. and Roy, R. 1996. *Continuum Percolation*. Cambridge University Press.

NI, J. AND CHANDLER, S. 1994. Connectivity properties of a random radio network. *IEE Proceedings-Communications. 141,* 4, 289–296.

OSHER, Y. 2007. From local search to global behavior: Ad hoc network example. *Cooperative Information Agents XI*, 196–208.

PENROSE, M. AND PISZTORA, A. 1996. Large deviations for discrete and continuous percolation. *Advances in Applied Probability. 28*, 29–52.

PHILIPS, T., PANWAR, S., AND TANTAWI, A. 1989. Connectivity properties of a packet radio network model. *IEEE Trans. on Information Theory. 35,* 5, 1044–1047.

PIRET, P. 1991. On the connectivity of radio networks. *IEEE Trans. on Information Theory. 37,* 5, 1490–1492.

SANTI, P. AND BLOUGH, D. 2003. The critical transmitting range for connectivity in sparse wireless ad hoc networks. *IEEE Trans. on Mobile Computing. 2,* 1, 25–39.

SANTI, P., BLOUGH, D., AND VAINSTEIN, F. 2001. A probabilistic analysis for the radio range assignment problem in ad hoc networks. In *Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing(MobiHoc)*. ACM.

TAKAGI, H. AND KLEINROCK, L. 1984. Optimal transmission ranges for randomly distributed packet radio terminals. *IEEE Trans. on Communications. 32,* 3, 246–257.

WAN, P. AND YI, C. 2004. Asymptotic critical transmission radius and critical neighbor number for k-connectivity in wireless ad hoc networks. In *Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing(MobiHoc)*. ACM.

XING, G., WANG, X., ZHANG, Y., LU, C., PLESS, R., AND GILL, C. 2005. Integrated coverage and connectivity configuration for energy conservation in sensor networks. *ACM Trans. on Sensor Networks. 1,* 1, 36–72.

XUE, F. AND KUMAR, P. R. 2004. The number of neighbors needed for connectivity of wireless networks. *Wireless Netowrks. 10,* 2, 169–181.

ZHANG, H. AND HOU, J. 2005. Maintaining sensing coverage and connectivity in large sensor networks. *Ad Hoc & Sensor Wireless Networks. 1,* 1-2, 89–124.