Analysis of Topological Properties of Random Wireless Sensor Networks

by Sergio Adolfo Bermudez-Rodriguez

ANALYSIS OF TOPOLOGICAL PROPERTIES OF RANDOM WIRELESS SENSOR NETWORKS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Sergio Adolfo Bermúdez Rodríguez

August 2010

ANALYSIS OF TOPOLOGICAL PROPERTIES OF RANDOM WIRELESS SENSOR NETWORKS

Sergio Adolfo Bermúdez Rodríguez, Ph.D.

Cornell University 2010

Connectivity of wireless networks has been widely studied. Within the framework of wireless sensor networks, asymptotic results have been provided for the probability that a network realization will be connected. However, exact formulas for finite network deployments are missing. This dissertation solves selected connectivity problems arising in wireless sensor networks of finite size. We approach the problem through stochastic geometry and combinatorial techniques.

The main topic considered in this dissertation is the connectivity of finite, randomly deployed, wireless sensor networks. The problem is to calculate the probability of connectivity given certain conditions like field of interest, number of node platforms deployed, type of infrastructure, and communication channel model. We provide analyses for multiple scenarios comprising one- and two-dimensional networks.

For one-dimensional network deployments, exact formulas for the probability of connectivity are given when deterministic communication links are considered along infrastructure. We also present an analysis of a network composed of nodes having random communication radii and provide a formula for general distribution functions.

For the two-dimensional network deployments approximate formulas for the probability of connectivity are provided. We consider deterministic and random communication links. In addition, we study the effects of partially connected wireless sensor network, where we allow

the existence of few isolated node platforms and thus can improve other network metrics, like node energy consumption.

Finally, we present an application of the obtained connectivity results. We focus on the extension of network functional lifetime through a topology control scheme. The scheme is based on the correlation of information obtained by the node platforms in the deployment. This analysis helps to illustrate the use and relevance of our results when designing wireless sensor networks.

BIOGRAPHICAL SKETCH

The author received the Bachelor of Science degree in Electronics and Communications Engineering from the Instituto Tecnológico y de Estudios Superiores de Monterrey (Monterrey Institute of Technology) in Monterrey, México, in December of 1999. Subsequently he worked in the related industry for several years. In August 2005 he became a graduate student in the School of Electrical and Computer Engineering at Cornell University and then a member of the Wireless Intelligent Systems Laboratory (WISL), under the guidance of Professor Stephen B. Wicker. His research interests include the modeling and analysis of connectivity, coverage, range assignment, energy consumption, and mobility of random wireless sensor networks.

a la suerte,

a las coincidencias,

al azar. . .

ACKNOWLEDGMENTS

I have had the privilege of interacting with bright and talented people. This stepping stone to become a researcher has been possible with the support of my advisor Professor Stephen B. Wicker. His optimism and guidance has been fundamental to complete this work. He made my graduate life in Cornell University an enjoyable experience. I express my gratitude to Professor A. Kevin Tang and Professor G. Edward Suh for being part of my special committee. I also praise the capacity and idealism of Professor Francisco J. Valero Cuevas, he has made the process of becoming a graduate student feasible for many people. A strong recognition to the Consejo Nacional de Ciencia y Tecnología (CONACYT) from México for sponsoring so many dreams.

Thanks to my friends and group members at Cornell University for their help and company. Special appreciation to Paola, Eurídice, Ivan, Matt, Onur, Javier, Rafael, Saswat, Sameer, Ania, Phil, Hui, Hazer, Christina, Coalton, Mikhail, and Daniel.

I am indebted to my old good friends and mentors, Liz, Israel, Alejandro, Delfino, Raúl, Andrés, Vincent, Juan M., and Mónica. It has been my pleasure and privilege to have coincided and lived our mutual experiences and time.

I am grateful to my parents and brothers for their love and support. Singular and exceptional credit to Olga for being, without her this idea would not have been possible, спасибо Ольга. . .

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1-D	One-dimensional object, *10*, *14*
2-D	Two-dimensional object, *10*
AWGN	Additive White Gaussian noise, *58*
BS	Base station or sink, *6*, *36*
CDF	Cumulative distribution function, *28*
FoI	Field of interest, *10*, *19*
iff	If and only if, *36*
i.i.d.	Independent and identically distributed, *25*
pdf	Probability density function, *30*
RF	Radio frequency, *21*
rv	Random variable, *17*
SHM	Structural Health Monitoring, *4*, *95*
wlog	Without loss of generality, *53*
wrt	With respect to, *35*
WSN	Wireless Sensor Network, *3*, *13*

LIST OF SYMBOLS

$\|\cdot\|$	Euclidean norm on $\mathbb{R}^n$, n from context, *36*
$[m]$	Set of integers $\{1, 2, \ldots, m\}$, *36*
$\mathbf{1}_{\{A\}}(\cdot)$	Indicator function over set A, *24*
$\Pr\{\cdot\}$	Probability measure over Euclidean space, *37*
$\perp\!\!\!\perp$	Probabilistic independence, *37*
$\mathbb{E}\{\cdot\}$	Expected value of a random variable, *30*
$\mathrm{B}(\cdot)$	Beta function, *40*
$\Gamma(\cdot)$	Gamma function, *28*
$\mathscr{N}(\mu, \sigma^2)$	Normal (Gaussian) distribution, mean μ, variance σ^2, *30*
$\Phi(\cdot)$	Distribution function of $\mathcal{N}(0, 1)$, *28*
$\mathcal{U}(A)$	Continuous uniform distribution over A, *38*
$F(X)$	Distribution function of random variable X, *28*
f_X	Probability density function of random variable X, *30*
$X \sim F$	Random variable X has a distribution function F, *22*
X_σ	$X \sim \mathcal{N}(0, \sigma^2)$, *60*
η	Path-loss exponent, *60*
Π_λ	Poisson spatial process with density λ, *52*
$N(A)$	Random variable, number of points of Π_λ in A, *53*
$\mathrm{b}(\boldsymbol{x}, r)$	Disc in $\mathbb{R}^2$ with center $\boldsymbol{x}$ and radius r, *56*
$\binom{n}{k}$	Binomial coefficient, n choose k, *38*

CHAPTER 1

INTRODUCTION

1.1 Topological Properties of Wireless Sensor Networks

A wireless sensor network (WSN) is composed of communicating nodes with sensing capabilities, called node platforms, [26, 43, 46, 65]. WSNs represent a special type of ad hoc wireless networks [36, 37]. In general, WSNs have large number of node platforms [65]. The use of great number of node platforms in a WSN demands a special type of analysis, different from traditional wireless networks [4, 46]. When deploying a large number of node platforms in a given environment, for a particular sensing application, it may be complex and expensive to perform a structured deployment [43, 46]. As a result, a random deployment of node platforms may be necessary in order to be able to use vast quantities of such platforms in a feasible and economical manner.

When designing a WSN to be comprised of a random deployments of node platforms, it is important to know the probability that the resulting wireless network will be connected. In particular, in order to provide certain guarantees that the resulting network will be connected with certain probability, it is necessary to estimate the number of node platforms that should be deployed. A network design engineer has the challenge to minimize this amount of nodes in order to provide a cost effective design.

In this dissertation we analyze some topological characteristics, as defined below, of WSNs. In particular we study the connectivity probability of WSNs with finite number of node platforms. Also, we present a simple topology control mechanism that helps to extend the network functional lifetime. Current methods researching wireless network connectivity issues focus on the probability of connectivity based on asymptotic results on the number of nodes in the network deployment. These asymptotic approaches provide estimates for the probability of network connectivity, as opposed to exact results, which are possible when analyzing a wireless network with finite number of nodes.

Once a network realization is connected and functional, it is important to estimate the energy consumption of the resulting topology with the goal of finding a way to improve the network functional lifetime of the given deployment. This dissertation proposes a simple topology control scheme that is capable of extending the network functional lifetime after the network has been established.

The practical relevance of this dissertation is as follows. In real world installations of WSNs the deployed quantities of node platforms are finite. Some application of such networks include monitoring physical structures—such as tunnels, bridges, or buildings, where WSNs are used to observe the mechanical conditions of the structures and to evaluate their possible damage [18]. As a result, in these types of environments it is relevant to have reliable estimates of the required finite number of node platforms to be deployed and take them into consideration when planning a network [28]. In addition, it is important to accurately model the communication channel, not only in a deterministic fashion

(Boolean-type communication links), but also in a random way, such as consideration of fading—for example log-normal shadowing, a technique that models obstructions.

Additionally, note that when node platforms are embedded into physical structures, like the ones mentioned above, having a wireless network functioning for long periods is of utmost relevance [52]. This follows given that it may be impossible to replace the source of energy of the node platforms—either because they are embedded or due to their sheer numbers. Even though there are some types of node platforms that are able to harvest energy [62], it is a desirable design approach to utilize the available energy in the most efficient way. A topology control mechanism is a viable and reasonable way to extend the network functional lifetime.

1.1.1 Background on WSNs

WSNs are a recent technology resulting from the miniaturization of both radio components and sensor devices [67]. A wireless sensor network (WSN) is composed of small devices with communication and sensing capabilities. These small devices are called node platforms, sensor platforms, motes, or simply nodes—in this work we will use any of these names interchangeably. These characteristics of WSNs allow for fine spatio-temporal sensing and monitoring of any given environment. But, on the other hand, they also create complex design challenges, including routing algorithms [46], energy consumption, time synchronization, mobility [51], and security and privacy [59].

WSNs have multiple applications, including the broad areas of military tracking, habitat monitoring, and patient health monitoring [50, 80, 86]. Some of these applications require a large number of node platforms to be deployed in the fields of interest in order to be able to obtain useful data while maintaining network robustness and architecture reliability. In particular, WSNs with high-densities of node platforms can be embedded into physical structures. For example WSNs can be used for the monitoring of buildings, bridges, and tunnels, for the purpose of performing techniques of Structural Health Monitoring (SHM). Currently, in new cutting-edge bridges and smart buildings there are wired SHM analyses being performed. It is of great interest to be able to perform such SHM analyses in a wireless fashion [58, 82].

In a general sense, network topology for traditional wired and wireless networks is defined by their deployment scheme. There are two wide category types of such deployments (1) structured—for example star or mesh topology—and (2) random—following a given distribution [46]. As a result of the kind of deployment, other characteristics of the networks are directly affected, for example connectivity [48, 76, 85] and coverage [1, 42, 49, 54, 55, 77], which in turn affect the capacity of the network deployment [39]. Thus the network connectivity and coverage can be considered topological characteristics of a network. Note that, given some conditions, connectivity can imply coverage in a wireless sensor network [3, 56, 79].

The topological characteristic of connectivity and coverage in wireless networks are crucial in WSNs given that certain design requirements should be met in order for a network to be able to perform its functions. In particular, connectivity is very relevant for wireless sensor networks because it allows a network user to be able to extract information out of the environment under sensing. Besides the number of node platforms being deployed, the characteristics of the environment where the platforms are deployed can greatly affect the level of connectivity of a WSN. Since WSNs should be robust and reliable when designed for real world applications, it is necessary to consider the type of channels where the nodes will be operating in order to be able to maintain the network connectivity even in harsh conditions.

In this dissertation we focus on the connectivity analysis of multi-hop, mesh-type topology, WSNs resulting from random deployments of node platforms. Note that when WSNs are used for monitoring physical structures, like in SHM, they have to be very energy efficient, given that node platforms will remain in place for long periods. Thus, after analyzing how to have a connected network, with a given probability, this work presents a mechanism that exploits the relation between the connectivity aspects and their impact on functional lifetime in a WSN.

1.1.2 Connectivity Problem in WSNs

The level of connectivity in general wireless networks is a relevant, intrinsic quality, since it allows the networks to fulfill their primary or elementary function of communication between the node platforms—

including sinks or base stations, if applicable—belonging to the network. Connectivity of a network can be broadly defined as the ability of any pair of nodes that are part of the network to be able to communicate with each other, either in single- or multi-hop mechanism. A network is fully connected—also simply called connected—if any pair of nodes is able to communicate between them. A network is partially connected if there are isolated nodes or components of nodes [22].

Having a connected network allows the transmission of data between nodes, or from a node to a single or multiple base stations (BSs), also called sinks. Among other things, data communication among network elements allows for data extraction, cooperative schemes, self-configuration protocols, as well as routing protocols establishment.

A connected network is easy to characterize: is the one where all the nodes can communicate to each other. However, besides this full connectivity concept, there can be a partially connected network. The partial network connectivity can be measured as the fraction of nodes that form a connected component. In the particular case when there is infrastructure, the partial connectivity could also be measured as the fraction of nodes that is able to reach a base station or sink.

In random wireless sensor networks only a probabilistic measure can be given regarding the connectivity degree of such networks. That is, the design goal of a random WSN is to provide certain guarantees for the connectivity of the resulting network. Among the different types of metrics for connectivity of random WSNs there are the following

- asymptotic connectivity
- infinite connected component
- connected subnetworks
- partial connectivity

Another factor to consider when analyzing the connectivity of a WSN is that the network can have infrastructure—like base stations or sinks—or the lack of it. Given the above mentioned characteristics for a particular deployment, there are techniques that improve the level of connectivity in wireless networks [13] by using multiple antennas in each node. This topic will not be considered in this dissertation.

In summary, the problem of connectivity in randomly deployed wireless sensor networks is to have a certain probabilistic measure that indicates the most likely connectivity degree of the resulting network realization. This guarantee will be a function of certain parameter of the network, like the node density or quantity of nodes in the deployment, and the area of the environment.

1.1.3 Energy Consumption Problem in WSNs

Improving the energy consumption in general wireless networks is important given that, most of the time, their nodes only have a finite energy source—like a battery. In particular, as mentioned above, having efficient energy consumption in WSNs is critical, due to the impossibility of replenishing the energy source in the node platforms. As a result, one of the most important goals of any design in WSNs is to have the most power efficient network possible. This limitation comprises all the aspects of the design work, including network protocols and architectures. In this dissertation we present a simple scheme that, after having a connected wireless sensor network, can be used to extend the network functional lifetime. This functional lifetime extension results from a reduction on energy consumption of the node platforms forming the wireless network.

Mostly, network lifetime in a WSN can be defined in several ways. The three most used definitions are when

- the first node failure occurs
- a given percentage of the number of nodes have failed
- the network cannot function as designed

The techniques of network lifetime extension are the bases for having long-lived networks. The methods apply specially when considering the monitoring of structures, as in SHM, where node platforms are usually embedded into the structure—thus with no possibility of having their energy sources replaced or recharged. A node lifetime extension can be directly measured in proportion of the duty cycle of the node platforms. Node duty cycle is the fraction of time that a node is in the *active* state, as opposed to the *sleeping* or *idle* state. Thus the lifetime extension metric for a node can be defined as the ratio between the expected lifetime of a node with a particular duty cycle and a node that is always active.

Broadly, there are different approaches to improve the network lifetime, each with its own characteristics, advantages, and disadvantages. Examples of approaches include (1) maximal, (2) maximization, and (3) extension of network lifetime. In this dissertation we focus on the last approach, that of extending the WSN lifetime through a simple scheme.

In summary, the problem of energy consumption in a WSN is related to that of extending its functional lifetime. In randomly deployed WSNs is it beneficial to have a mechanism that, once the network is connected, will allow an increase to the total time the network will be functional, compared to a deployment without such a scheme.

1.1.4 Stochastic Analysis of WSNs

Throughout this dissertation we use the theory of applied probability [29,30], stochastic processes [44,64,66], and random graphs [15,61] to address the problems mentioned above. These techniques are useful to analyze random wireless sensor networks [73, 74]. In particular, by mapping a network deployment to the abstract mathematical concept of a graph it is possible to rigorously analyze a random WSN. This abstract mapping is based on the following equivalences:

- node platforms are represented by vertices,
- communication links become edges, and
- conditions on the communication channel provide the restrictions for the existence of edges

As a result of this network-graph mapping, the field of interest (FoI) of the network deployment becomes one of the Euclidean spaces. Both the one-dimensional (1-D) and the two-dimensional (2-D) Euclidean spaces are the ones of interest in this work. We note that an alternate approach to analyze the problems considered in this dissertation is via stochastic simulations [2]. One of the most relevant disadvantages of this approach, as compared to the analytical method, is the scalability issues on the computational simulations [81].

Using stochastic simulations to obtain the probability of connectivity is a resource-intensive computational method, given that the algorithms to obtain such probability require extensive search for verifying the connectivity of each network realization. In addition, when using the Monte Carlo method for simulations it is necessary to provide enough replications of the same simulation in order to obtain the appropriate level of confidence intervals for each set of parameters. In this work we use stochastic simulations for small-sized—in the number of nodes in a WSN—problems to verify our theoretical results.

1.2 Dissertation Outline and Summary of Results

This dissertation addresses some of the problems of connectivity in one and two dimensional random wireless sensor networks. In addition, this work presents a simple scheme that increases the network functional lifetime and it is based on the connectivity results from the two-dimensional WSNs. Our main contributions are the formulas for the probability of network connectivity and network functional lifetime extension. In particular, this dissertation:

- provides an unifying framework on the probability of connectivity for finite random WSNs, considering 1-D and 2-D deployments;
- develops new methodology to analyze the probability of connectivity of finite random wireless sensor networks;
- develops new methodology to analyze the probability of connectivity of WSN with infrastructure;
- explores the concept of partial connectivity in WSNs with infrastructure;
- shows a simple mechanism that controls the network topology with the purpose of extending its functional lifetime.

The outline of this dissertation is as follows.

In Chapter 2 we address the network connectivity in one-dimensional WSNs deployments. We analyze networks with a finite number of node platforms under different conditions. These settings include deployments of nodes with non-homogeneous (random) transmission capabilities and networks with infrastructure.

In Chapter 3 we analyze network connectivity in two-dimensional WSNs deployments. We focus on networks with finite number of nodes and consider deployments with and without infrastructure. The analysis of the latter one includes concepts of partial connectivity.

In Chapter 4 we present a simple topology control scheme that extends the network functional lifetime on two-dimensional WSNs deployments. That Chapter illustrates the use of the techniques and formulas developed in the previous Chapters.

Chapter 5 concludes the dissertation with a brief summary of our contributions, an enumeration of open questions, and possible ideas of extensions for this work.

1.2.1 Related Publications

Some portions of this dissertation have been previously published in the research in [5–9]. In [5, 8] the probability of network connectivity under different settings is analyzed for a 1-D WSN model. In [6, 7] the probability of connectivity for a 2-D WSN model is studied, along with a simple topology control mechanism. Finally, [9] presents the analysis of full and partial network connectivity for a 2-D wireless sensor network model with infrastructure.

CHAPTER 2

CONNECTIVITY ANALYSES IN 1-D WIRELESS SENSOR NETWORKS

2.1 Introduction

In this Chapter we consider the connectivity problem for networks of wireless monitoring devices on highly linear structures such as bridges, tunnels, and walkways. Connectivity is a fundamental characteristic of general wireless networks since it is required in order to be able to extract useful information out of such network. In particular, if we consider wireless sensor networks (WSNs) and their multiple applications, then connectivity is essential for the proper function of such networks.

Given that WSNs can monitor several kinds of environments it is important to be able to estimate the level of connectivity of a random network realization at the time of its design. For example, WSNs can be used for monitoring physical structures such as bridges, tunnels, and walkways. This is comprised by structural health monitoring (SHM). It is envisioned that nodes will be embedded into the structures at the time of their construction. We note that WSNs have multiple design challenges, comprising various spheres, like network topology, node energy consumption, computing processing capabilities, and communication reliability and robustness.

In this Chapter we focus on WSNs formed by a random deployment of node platforms and analyze the probability that such network realizations will be connected. By analyzing a simple environment we are able to obtain exact formulas. In particular, in this Chapter we center in one-dimensional (1-D) wireless networks with and without infrastructure, where finite number of nodes are randomly and uniformly deployed.

First, when considering 1-D WSNs with infrastructure, we suppose that node platforms have deterministic communication links and provide exact formulas for their probability of connectivity. Then, when considering 1-D wireless networks without infrastructure, we assume a model where node platforms have random communication range with distribution function F and provide a formula for the probability of network connectivity for a general function F. In addition, we provide a comparison of such wireless network model and the result of a theoretic setting that assumes slow fading in the communication channel—by modeling it as a log-normal shadowing.

We note that in the research literature the problem of estimating the probability of network connectivity has been addressed for 1-D finite wireless networks. Although with the assumptions of infrastructure-less networks and node platforms with deterministic communication links [21, 35]. In this Chapter we consider WSNs with infrastructure as well as the situation of a wireless network composed of nodes having random communication radii.

The techniques we use to solve the connectivity problem presented in this Chapter are stochastic geometry and combinatorial theory. Our main contributions are the general formulas for the probability of network connectivity. In particular, we provide exact formulas for both cases of node communication radii: (1) homogeneous and (2) random.

Moreover, we also show how a network model that assumes node platforms with random communication radii following the uniform distribution function can be used to approximate the setting of a WSN being analyzed under a communication channel modeled with log-normal shadowing.

In addition, in this Chapter we also analyze a 1-D wireless network with infrastructure, that is with base stations. By using combinatorial theory we provide formulas to calculate the probability that such a network will be connected for n uniformly deployed nodes and m base stations with given locations.

This Chapter is organized as follows. Next Section 2.2 presents the related work in the research literature to this Chapter. Then in §2.3 we present the results of probability of network connectivity in 1-D WSNs, first by addressing the setting of networks with nodes having Boolean, homogeneous communication links. Subsequently, we address the situation of WSNs with nodes having random communication radii. This model is useful given that is can approximate a WSN under a communication channel affected by fading, in particular by log-normal shadowing model. In §2.4 we study the situation when the network deployment has several sinks or base station—that is a WSN with infrastructure—while assuming nodes having Boolean communication links.

2.2 Related Work

The following research work is related to the topics in this Chapter. We group the literature review according to the subject referred to in each paper. Let us first consider research that uses a wireless network model with nodes having homogeneous and Boolean communication radii or links. The following papers consider infinite size network models; where infinite size is in either the area of the field of interest or the number of nodes in the deployment. The percolation of a message—the distribution of the distance traveled by a broadcast sent by a source—is analyzed in [17]. The authors model random wireless network and find that no percolation is possible.

In [38] there is a study on the connectivity of wireless networks with n uniformly distributed nodes over a finite area. That work considers the homogeneous communication radius required for asymptotic connectivity when n increases. The authors use percolation theory to prove results in dense networks—networks with finite area, but infinite number of nodes. The work in [11] analyzes the connectivity of ad hoc networks. The author uses random geometric graphs theory and by estimating the probability that there is no node isolated in an infinite network with uniform density, he provides a formula that approximates—a lower bound—the probability of network connectivity. The authors in [25] address the connectivity in ad hoc networks by using percolation theory. That paper assumes an infinite size network and provides the requirements in the density of nodes to have an infinite connected component with high probability.

The following related work considers the connectivity of 1-D wireless networks with a finite number of randomly deployed nodes within a finite field of interest. Note that none of them takes into account neither infrastructure nor the randomness of the communication channel. The work in [35] treats random interval graphs considering in particular the unit interval and analyzing the random variables (rv) related to graph connectivity. The authors use combinatorial theory and provide several results including the network connectivity probability.

The authors in [21] analyze a network deployed over a line segment with n nodes distributed uniformly. Using the ratio of the volume of a convex polytope and an n-dimensional simplex, the paper provides a formula for the probability of having a connected network. By means of order statistics principles [33] provides a formula for the probability that a wireless network is composed of at most D clusters of connected nodes—note that a cluster can be just an isolated node.

The subsequent research work considers randomness in the communication channel, therefore these models have more complex node communication radii representation, those where there are a deterministic and a random part—or, equivalently, there are a distance based part and a channel-dependent part. The authors in [16] use percolation theory to analyze infinite wireless networks and their results hold when nodes have random communication radii.

In [57] there is an analysis of a 1-D wireless network over a semi-infinite line. Nodes are deployed following either a Poisson process or a distribution with general inter-node distances. Using queueing theory the authors provide formulas for the probability that a node at given distance will be connected to a base station at the origin. The work considers both a deterministic channel model and one assuming fading and log-normal shadowing, both in infinite networks.

In [12] there is an analysis of the network connectivity under log-normal shadowing model where nodes are deployed over an infinite plane according to a Poisson process. The work provides bounds on the probability of connectivity utilizing approximations based on geometric random graphs. The authors find out an increase in the network connectivity due to shadowing. It is also worth to note that under [12] analysis it is assumed statistical decoupling between the path-loss exponent and the variance in the log-normal shadowing model—something that may not happen in real world settings.

2.3 WSNs with Random Communication Range Nodes

As mentioned before, one of the most important topological characteristics of a wireless network is its connectivity. It is a basic quality of a WSN since it is required to obtain useful information out of such network. When designing a WSN, one of the factors that needs to be addressed is how many nodes to deploy. This Section analyzes the connectivity of WSNs considering finite number of randomly deployed nodes with random communication ranges—in contrast to a homogeneous communication range model as in most of the literature. In particular, it focuses on the probability of being able to convey a message from the source to

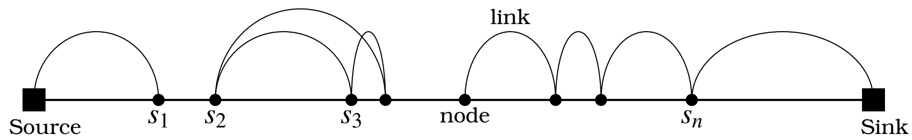

Figure 2.1: Example showing a realization of a network. Note that due to the effects of random communication radii node s_2 and s_3 are connected, but s_1 and s_2 are not in spite of their relative distance between each other.

the sink given certain number of nodes; a so called relay network. The setup is as follows. There is one source and one sink at the extremes of the field of interest and there are a given number of nodes, deployed randomly, between them. Then we need to estimate the probability that a message sent from the source is able to get to the sink.

To analyze the connectivity with random communication ranges we consider a one-dimensional (1-D) wireless network. The method to address the connectivity problem is by mapping it to the question of covering a circle with arcs of random sizes. This Section provides an analytical solution when the communication ranges of the nodes are distributed uniformly over the field of interest (FoI). See Fig. 2.1 for an illustration. We present formulas to calculate the probability of having connected network given a finite number of node platforms.

In addition, and of practical relevance, this Section shows how the random communication radii model can represent certain kind of randomness in the communication channel. Specifically there is a comparison between our model and the log-normal shadowing statistical model. Stochastic simulations validate the derived formula as well as its approximation to the standard log-normal shadowing model. In agreement with previous results for wireless networks, we verify that the requirements on the number of nodes to have a connected network may be relaxed when the nodes have random communication radii as compared to homogeneous transmission radius nodes.

There are several results in the research literature on wireless networks that analyze connectivity [5, 12, 21, 35, 38, 57]. In the case of analysis under asymptotic behavior, some of those works consider randomness in the communication channel by assuming the effects of log-normal shadowing—also called slow fading. On the other hand, the results on connectivity with finite number of nodes do not consider the randomness in the communication channel. In this latter case, the research models focus only on a Boolean—or binary—communication link model and nodes having homogeneous communication radii.

The random range model is relevant for network design because it can account for obstructions or variability in the communication links or channels, for example to be able to accurately estimate the probability of connectivity after network setup. Moreover, sometimes it is not reasonable to assume a homogeneous communication radius for all the nodes in the network, even worse, the network could use heterogeneous nodes. For instance, in the framework of monitoring physical infrastructure with wireless sensor networks (WSNs), a model that takes into consideration channel randomness is important to have a more realistic representation. This follows because nodes will be embedded into the structures—homes, buildings, tunnels, bridges—thus having links with arbitrary obstructions. The random range model is also applicable on networks deployed after an emergency in a urban or combat search and rescue operation, where a source needs to convey an urgent message to a destination or control center and there are unplanned obstructions.

By assuming that the communication radii of the nodes will be an independent and identically distributed uniform random variable with finite support, the model in this Section can represent a kind of randomness in the communication channel. Under certain restrictions, the model could be used as an approximation to a shadowing model. Log-normal shadowing model is a widely accepted statistical model used to analyze the randomness in communication channels. This model represents the slow fading of the radio frequency (RF) communication signals and, with appropriate parameters, it is valid for either indoor or outdoor environments.

Therefore, with some limitations, our model represents a simple approach to deal analytically with the complexity of log-normal shadowing (referred simply as shadowing below). As noted in the research work that study network connectivity, shadowing, or randomness in the communication channel, increases the probability of having a connected network. The 1-D setup of this Section also exhibits this phenomenon, that is there are slightly more relaxed requirements on the number of nodes needed to have a connected network, compared to those needed for a connected network using a Boolean model with homogeneous radius of communication.

2.3.1 Model and Problem Formulation

Let the segment $\mathcal{I} = (0, 1)$ represent the field of interest. Consider a deployment of n nodes over $\mathcal{I}$ following a uniform distribution. Label the nodes $s_1, s_2, \ldots, s_n$ in increasing order according to their location, starting from the leftmost one, and denote by x_i the position of node s_i. Assume there is a source s_0 at $x_0 = 0$ and a sink s_{n+1} at $x_{n+1} = 1$. The source wants to send a message to the sink. A similar model, with a semi-infinite line, has been used in the literature for the *percolation* of a message [17, 57].

Consider that node i has a random transmission power such that its communication radius r_i is a uniform random variable, $r_i \sim \mathcal{U}(0, 1)$, $i \in \{0, 1, 2, \ldots, n\}$. That is when node s_i, $i \in \{0, 1, 2, \ldots, n\}$ is deployed it is assigned an independent random variable r_i. Having the source and the sink, we want to find the probability that if we deploy n nodes over the segment $\mathcal{I}$ the resulting network is connected.

The procedure to solve this problem is by first mapping it as a circle covering problem, then solving the covering problem, and finally interpreting the result. Fig. 2.1 illustrates a realization of a network where squares are the base stations, circles are the nodes, and arcs denote communication links. Note that due to the effect of random communication radii node s_2 and s_3 are connected, but not s_1 and s_2 in spite of their relative distance between each other. Fig. 2.2 is an equivalent representation of the network, where the vertical side of a triangle represents the communication range, or radius, of a node. Since we are interested in relaying a message from source to sink it suffices to analyze the one-sided connectivity of this network. Then a relay link between a

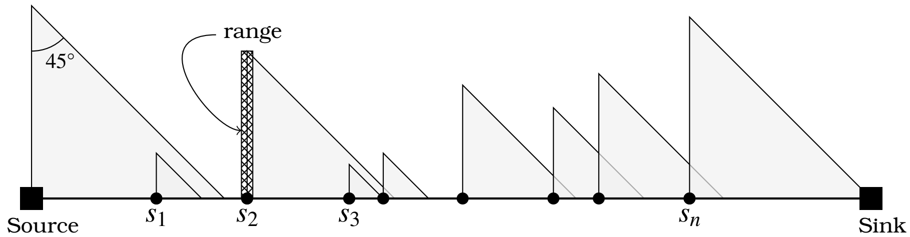

Figure 2.2: Equivalent network representation of Fig. 2.1. A communication link, in source to sink direction, between a pair of nodes exists if the triangle originated in the leftmost node covers the rightmost node.

pair of nodes exists if the triangle originated in the leftmost node covers the rightmost one.

The problem statement is the following:

Let the source be at $x_0 = 0$, the sink at $x_{n+1} = 1$, and n nodes uniformly distributed over $\mathcal{I}$. Let node i have a random communication radius $r_i \sim \mathcal{U}(0, 1)$. Denote by C_n the event that a message sent from the source will reach the sink. Then what is the probability of the event C_n?

2.3.2 Analysis

Homogeneous Communication Radii

First let us consider the connectivity of a WSN with nodes having homogeneous communication radii $r_c = r$. This result is used to establish a comparison against the event C_n and to obtain an insight for the proof of the main result of this Section. In this setup, node s_i at location x_i

transmits and node s_j at location x_j can receive the transmission if and only if their distance is less or equal than r, that is $\|x_i - x_j\| \leq r$. Where $\|\cdot\|$ is the Euclidean norm.

Let $C_{n,r}$ be the event that a network with a source at $x_0 = 0$, a sink at $x_{n+1} = 1$, and n nodes uniformly deployed over $\mathcal{I}$, each with $r_c = r$, is connected. Then as shown below in §2.4

$$\Pr\{C_{n,r}\} = \sum_{j=0}^{n} \binom{n}{j} (-1)^j (1 - jr)^{n-1} \mathbf{1}_{\{1 \geq jr\}} \tag{2.1}$$

where $\mathbf{1}_{\{\cdot\}}$ represents the indicator function

$$\mathbf{1}_{\{w \geq r\}} = \begin{cases} 1, & \text{if } w \geq r; \\ 0, & \text{otherwise.} \end{cases}$$

Now we mention an insight that will be useful for a proof below. Note that the notion of network connectivity using n nodes with homogeneous communication radii and two base stations—that is $n + 2$ total number of points—over $\mathcal{I}$ is equivalent to covering a circle having unit circumference with $n + 1$ uniformly positioned arcs of equal length. The circle circumference can be scaled by $\|\mathcal{I}\|$ if necessary. This circle covering problem was solved in [72]. The equivalence follows by representing a node as the counterclockwise border of an arc and the length of the arc is the size of r_c. Considering homogeneous r_c implies that links are symmetric, then two nodes are connected if and only if their corresponding arcs in the circle intersect.

Random Communication Radii

Now consider that nodes have a random transmission power such that their communication radii r_c is a uniform random variable over $\mathcal{I}$. See Fig. 2.2 for a representation. As we mentioned, this means that when node s_i, $i \in \{0, 1, 2, \ldots, n\}$ is deployed it is assigned an independent random variable r_i. When the communication radius is homogeneous the result is provided by (2.1). We want to obtain conditions on the number of nodes that allow a message sent from the source to reach the sink in a multi-hop fashion.

Theorem 2.1 (Uniform r_c). *Assume that node s_i has communication radius r_i, $i \in \{0, 1, \ldots, n\}$, where r_i's are independent and identically distributed (i.i.d.) random variables with uniform distribution $\mathcal{U}(0, 1)$ and r_0 has the largest radius of the realization, then the probability of the event that this network is connected, C_n, is given by*

$$\Pr\{C_n\} = 1 + \frac{n!}{(2n-1)!} \sum_{k=1}^{n} \left\{ \frac{(-1)^k}{(n-k)!\, k\, 2^{n-k}} \sum_{\mathcal{S}} \left[\binom{n-k}{m_1, \ldots, m_k} \prod_{i=1}^{k} (2m_i + 1)! \right] \right\} \tag{2.2}$$

the inner sum is over all sets in $\mathcal{S}$, where $\mathcal{S}$ is the collection of all k-element sets with nonnegative integers m_i that satisfy

$$\sum_{i=1}^{k} m_i = n - k$$

Proof. Let us construct a mapping between network connectivity and the covering of a unit circumference using randomly located arcs of random sizes. This technique facilitates the connectivity analysis by eliminating the boundary conditions on the network deployment.

First, consider a network over $\mathcal{I}$ with node s_i having communication radius r_i, located at position x_i, $i \in \{0, 1, \ldots, n\}$ and r_0 being the largest radius. Then, on a unit circumference, take r_i to be the length of the arc i and mark as s_i the counterclockwise extreme of this arc. As a reference, mark the place in the circumference where s_0 is located as o. Place s_1 on the circumference at the corresponding distance x_1 along the circumference, measured from o. Continue deploying the rest of the nodes s_i together with their corresponding arcs i with length r_i. Set s_{n+1} also at o. This completes the mapping of the network onto the circle.

Now, assume a realization of the covering of a circumference with arcs of random sizes. Select the longest arc and mark its counterclockwise extreme as o, this point represents the source s_0 on the network. Going clockwise, assign the length of arc i to the communications radius r_i of node s_i and the length along the circumference from o to the counterclockwise extreme of arc i to x_i in the network, $i \in \{0, 1, \ldots, n\}$. Set $x_{n+1} = 1$. This completes the mapping of the circle onto the network. Fig. 2.3 is an equivalent representation of Fig. 2.2, thus of Fig. 2.1.

As a result of the previous mapping, the unit circumference is covered if and only if the network over $\mathcal{I}$ is connected. Hence our problem reduces to that of covering a circle with arcs of random sizes. We note a caveat, the procedure to solve for C_n by mapping network connectivity to circle covering is valid only when the largest r_i is r_0. This could be justified if we consider the source s_0 to be the node with more resources

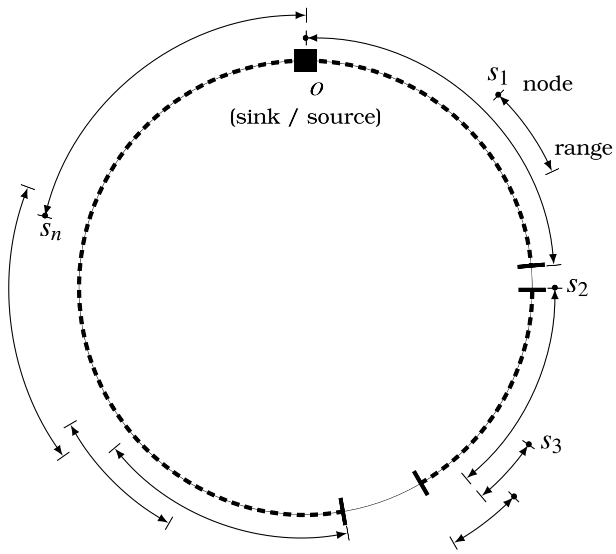

Figure 2.3: Representation of the connectivity of the network realization in Fig. 2.2 by the covering of a circle with unitary circumference.

on the network deployment. The fact that r_0 is the largest r_i assures that, on the circle, the arc n will not cover s_1—this would imply a link between s_n and s_1 in the network.

To analyze the covering of a circle we use the results by Siegel and Holst [71]. Consider n arcs placed uniformly and independently on a circle, each of a random size. The authors provide a formula for the distribution of the number of uncovered gaps N_n on the circle, that is $\Pr\{N_n = m\}$, when arc sizes are i.i.d. random variables with distribution function F over $\mathcal{I} = (0, 1)$. In our problem, to completely cover the circle set $m = 0$, then the event $\{N_n = 0\}$ equals $\{C_n\}$. Now when $m = 0$ the general formula for $\Pr\{N_n = m\}$ in [71] simplifies to

$$\Pr\{C_n\} = \sum_{k=0}^{n} \left\{ (-1)^k \binom{n}{k} (k-1)! \int_{\mathcal{I}} \Big[\prod_{i=1}^{k} F(u_i) \Big] \Big[\sum_{j=1}^{k} \int_0^{u_j} F(v)\,\mathrm{d}v \Big]^{n-k} \mathrm{d}u \right\} \tag{2.3}$$

where $\mathcal{T}$ denotes the collection of k-element sets of spacings u_i between k independent uniform points on the circle. Since those k points represent the location of the nodes, then by construction we have

$$\sum_{i=1}^{k} u_i = 1$$

Specializing (2.3) for the distribution function or cumulative distribution function (CDF) $F(x) = x^\alpha$, $0 < x < 1$, $\alpha > 0$, the formula simplifies

$$\Pr\{C_n\} = \sum_{k=0}^{n} \Biggl\{ (-1)^k \binom{n}{k} \frac{(k-1)!}{(\alpha+1)^{n-k}\,\Gamma((\alpha+1)n)} \cdot \sum_{\mathcal{S}} \Biggl[\binom{n-k}{m_1, \ldots, m_k} \prod_{i=1}^{k} \Gamma((\alpha+1)(m_i+1)) \Biggr] \Biggr\} \tag{2.4}$$

the inner sum is over all sets in $\mathcal{S}$, where $\mathcal{S}$ is the collection of all k-element sets with nonnegative integers m_i that satisfy

$$\sum_{i=1}^{k} m_i = n - k$$

$\Gamma(x)$ is the Gamma function, defined as

$$\Gamma(x) = \int_0^\infty (t)^{x-1} e^{-t}\, dt$$

where if x is a positive integer, then $\Gamma(x) = (x-1)!$. If $\alpha = 1$, then $F(x) = x$, the distribution function of a uniform random variable on $(0, 1)$.

Hence let $\alpha = 1$ in (2.4); for $k = 0$, $\Pr\{C_0\}$ reduces to $\Pr\{N_0 = 0\} = 1$; for $k > 0$ the terms with the Gamma function become $\Gamma(2n) = (2n-1)!$ and $\Gamma(2(m_i+1)) = (2m_i+1)!$. Thus the result follows. □

Corollary 2.1 (Random r_c). *If r_i's are i.i.d. random variables with general distribution function F with support on $(0,1)$, then the probability of network connectivity is given by* (2.3).

Note that by using (2.3) it is possible to define any suitable distribution F that describes best the application under consideration and obtain analytical results on the probability of connectivity.

Log-Normal Shadowing Approximation

A wireless network with nodes having random communication radii could serve as a model of a comparable network under log-normal shadowing. In a wireless channel there is RF signal attenuation that can be modeled by a large-scale path-loss with parameter η, the path-loss exponent. In addition, and due to obstructions in the communication links, there is a random attenuation of the RF signal. A well accepted statistical model that accounts for these random variations—in both indoors and outdoors environments—is the log-normal shadowing model. The shadowing model has as parameter σ, the standard deviation of the underlying random variable. The formula for the total path-loss, L, in decibels (dB), for a given distance, d, in meters (m), between transmitter and receiver, and for a channel model with parameters η and σ is

$$L(d,\sigma) = \bar{L}(d_0) + 10\eta\log\left(\frac{d}{d_0}\right) + X_\sigma \qquad \text{[dB]} \tag{2.5}$$

where $d_0 \ll d$ [m] is the reference distance where the reference power $\bar{L}(d_0)$ [dB] is measured. X is a zero mean normally distributed random variable with variance σ [dB], $\mathcal{N}(0, \sigma^2)$. Therefore L is also a random variable. In practical environments the values of σ fall within $(1, 10)$, while those of η range within $(2, 6)$ [63].

When considering the shadowing model, a node located at a certain distance from the transmitter would be able to correctly receive data if the received power of the RF signal is above certain threshold, call it β_{th}. The random r_c model from the previous Subsection can approximate a shadowing model in the following situation. When $r_c \sim \mathcal{U}(0, 1)$, its average over $\mathcal{I}$ is $\mathbb{E}\{r_c\} = \bar{r} = 0.5$. Where $\mathbb{E}\{\cdot\}$ denotes the expected value of a random variable. Then for a given η let us assume that β_{th} is such that when $\sigma = 0$, that is no shadowing, the average communication distance between a pair of nodes is then 0.5.

Note that X has infinite support. Since the field of interest is $\mathcal{I} = (0, 1)$, then for practical considerations the random variations induced in (2.5) by X should be confined in such a way that all the resulting communication links due to the path-loss L fall within $\mathcal{I}$. Denote by Y the random variable resulting from truncating X to the finite support $(-a, a)$ that achieves $r_c \in [0, 1]$. By the properties of the normal distribution, when σ increases the Section of Y under consideration—its support—becomes *flatter*. Then the probability density function (pdf) of Y tends toward the density function of a uniform random variable. As a result, when σ increases, r_c behaves like a uniform random variable.

Table 2.1: Parameters used in the simulation setup.

Parameter	**Value**
reps	10^4
n	$\{1, 2, \ldots, 12\}$
$\mathcal{I}$	$[0, 1]$
$\bar{r}$	0.5
r_c	$\mathcal{U}(0, 1)$
η	4
σ	0

2.3.3 Simulation Results

This Section presents simulation results that validate the developed analytic formulas for the network model in the previous Section 2.3.2. In addition it exhibits the comparison of this network scheme versus one considering shadowing in the communication channel. The simulations use the Monte Carlo method with 10^4 random replications for each set of parameters. Table 2.1 provides the simulation parameters.

The simulation technique is as follows. Take a deployment with n nodes and verify if the resulting network is connected. Repeat the procedure for the required number of replications for each n of interest. To calculate the probability of connectivity for a given n, take the ratio of the resulting total number of connected networks and all the network realizations for the corresponding n. For clarity purposes the confidence intervals are not depicted in the resulting graphs, though we remark that they were small—the maximum deviation width was less than 0.9% of the estimated average value for the probability of connectivity.

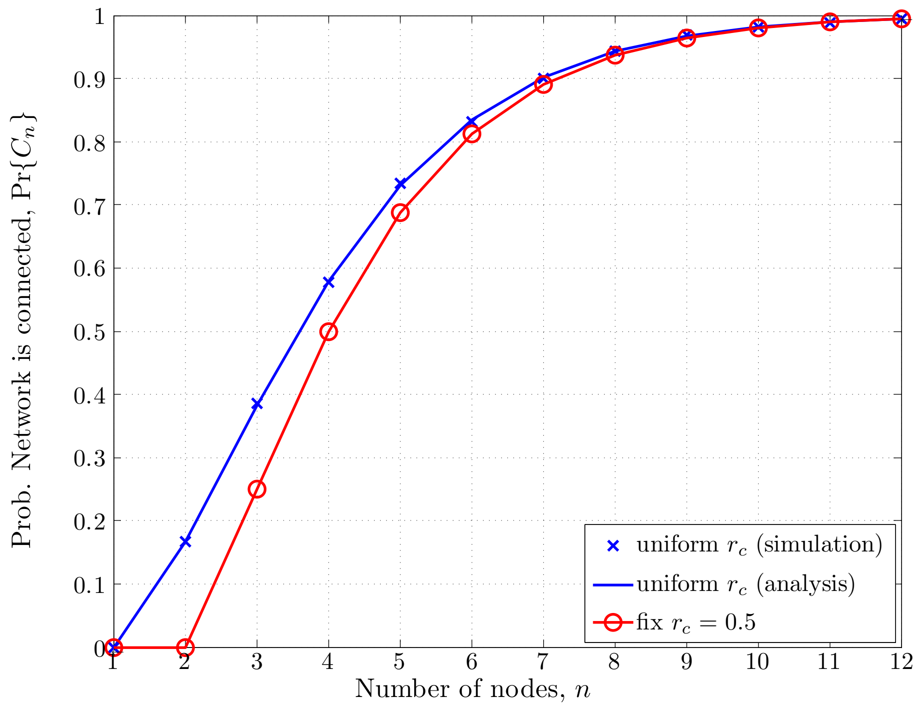

Figure 2.4: Comparison of the probability of having a connected network with respect to the number of nodes deployed between randomly uniform and homogeneous communication radii.

Results for Random Communication Radii

Using (2.2) we calculate the probability of connectivity for different n when $r_c \sim \mathcal{U}(0, 1)$. For comparison, using (2.1) we also calculate the probability of connectivity for a network with nodes having $r_c = \bar{r}$. Fig. 2.4 shows the simulation results. Solid lines represent the probability of having a connected network with respect to the number of nodes deployed. The line with crosses represents a network with nodes having uniform communication radii while the line with circles represents a network with nodes having fix r_c. For small n, note the slight increase in the probability of connectivity when considering nodes having uniform communication radii.

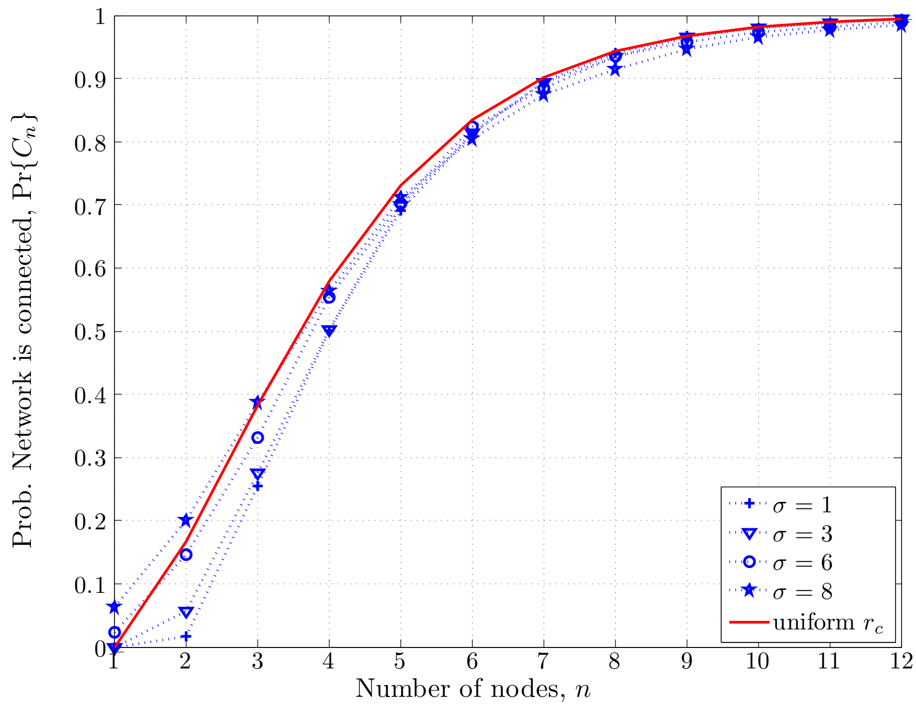

Figure 2.5: Comparison of the probability of having a connected network with respect to the number of nodes deployed between randomly uniform communication radius and a network under log-normal shadowing model, under different σ.

Results for Log-Normal Shadowing Approximation

Now consider a setting of practical relevance, a network under shadowing model. For the parameters of the shadowing model described in §2.3.2 take the average homogeneous r_c to be 0.5. Let the values of the standard deviation σ range between 1 and 8 and the path-loss $\eta = 4$. Without loss of generality, define β_{th} to be such that $r_c = 0.5$ when there is no shadowing, i.e. $\sigma = 0$. Fig. 2.5 shows the simulation results. Lines represent the probability of having a connected network with respect to the number of nodes deployed. The solid line represents the network with uniform communication radii, while the lines with markers represent the network under shadowing model for the different values of σ. Note that when σ increases the approximation is better. Large values of σ decreases connectivity, in agreement to previous results [12].

Results Discussion

First note that the analytic formula and the simulation agree. The slight increase in the probability of connectivity when considering nodes with uniform r_c in §2.3.2 holds only when n is small. This is due to the variations of r_c. When such radius is sampled from $\mathcal{U}(0, 1)$ it is possible for it to cover more length than $\bar{r}$, therefore resulting in a connected network and this effect is more prominent for small n. As a result of those variations in length, the probability to reach a second-hop node increases, thus the total probability of connectivity increases.

On the contrary, when n increases, those variations on r_c do not play a significant role because in this situation the node degree for most nodes is larger than 1. This effect is in agreement with that for 2-D networks [12], and for 1-D infinite networks [57]. As it can be seen in the results, when using a model that considers nodes with random communication radii, or assumes log-normal shadowing in the channel, simulations will show that it is possible to have a connected network with fewer nodes than those predicted by a model that does not take into account the randomness. Although, note that in real RF channels the parameters η and σ of the log-normal shadowing random variable are not independent, as it was assumed through this Section. In general, a higher standard deviation σ will imply a larger path-loss η.

2.4 Infrastructure Based WSNs under a Deterministic Communication Channel

In this Section we consider wireless sensor networks with infrastructure. Those networks consist of a finite number of randomly deployed node platforms and a backbone of access points at given locations. We provide an exact expression for the probability of having a fully connected network—every randomly deployed platform can reach at least one of the access points through either a single- or multi-hop path.

As mentioned above in §2.2, the research literature has focused on the connectivity of wireless networks having a single connected component, typically taking the form of a set of sensor platforms in which there is a connected path between any pair of sensors. In this Section we consider a variation of this problem in which a series of access points—also called sinks or base stations—are deployed along a linear structure such as a bridge or a tunnel. We derive a combinatorial expression for the likelihood of complete connectivity for a set of wireless sensors randomly distributed along this line—each sensor is to have connectivity with at least one sink. This physical setting may arise, for example, if wireless sensors are placed in the building materials prior to construction of homes, office buildings, or roads. We note that the presence of infrastructure—the base stations—significantly increases the likelihood of connectivity with respect to (wrt) a purely ad-hoc network, that is with no infrastructure. This is of strong relevance when designing a WSN.

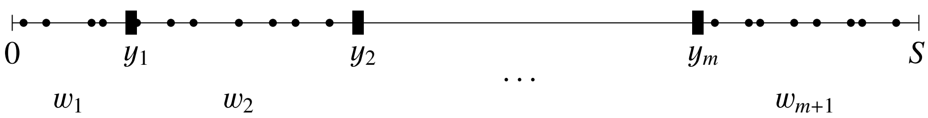

Figure 2.6: Network model over $[0, S]$; y_i represents a sink, w_i inter-sink points set, and dots over the line are wireless nodes.

2.4.1 Model and Problem Formulation

Consider a random deployment of n nodes over a line segment $\mathcal{I} = [0, S]$ following a uniform distribution. Label the nodes in ascending order starting from the leftmost one and denote by x_i the position of node i. Assume there are m sinks or base stations (BS) located at points $y_i \in \mathcal{I}$, $i \in [m]$; where $[m]$ represents the set of integers $\{1, 2, \ldots, m\}$. Further assume the set of base stations forms a backbone, this may be either wired or wireless. Let w_j denote the line segment $[y_{j-1}, y_j]$, with $j \in [m+1]$, and define $y_0 = 0$, $y_{m+1} = S$. Let k_j be the number of nodes in segment w_j. Fig. 2.6 illustrates the network abstraction of this model.

Definition 2.1. *Connected Subnetwork is defined to be the connected components of the network realization that are able to communicate with at least one sink.*

A wireless network is said to be connected if all its constituent nodes are part of a connected subnetwork. Consider a fixed communication radius r for every node. We assume that a pair of nodes share a communication link if and only if (iff) they are separated by a distance less or equal to r, equivalently $\|x_i - x_j\| \leq r$. Where $\|\cdot\|$ is the Euclidean norm. Note that, as a result of this model formulation, all the links are symmetric.

The problem statement is as follows:

> *Given n nodes with communication radii r, and m sinks and their locations, what is the probability that the resulting network is connected?*

2.4.2 Analysis

We solve the problem by decomposing it into independent subproblems. If we consider the interval $\mathcal{I}$ and a segment $w_j \subseteq \mathcal{I},\ j \in [m+1]$, then conditioned on the number of nodes k_j being in w_j the distribution of nodes within that segment follows again a uniform distribution. And each segment is probabilistically independent ($\perp$) of each other. There are only two particular cases to consider, according to the subsegment w_j where the nodes are deployed. These are:

1. Border-connectivity: there is only one sink at any one border of the segment w_i, $i \in \{1, m+1\}$.
2. Inner-connectivity: two sinks, one at each extreme of the segment w_i, $i \in \{2, 3, \ldots, m\}$.

Fig. 2.7 illustrates the decomposition principle.

Let C be the event when all nodes in the network reach at least one sink. Let C_i be the event that all nodes inside segment w_i are able to reach at least one sink. Set $\Pr\{C_i\} = 1$ if $k_i = 0$, then by the properties of the uniform deployment of nodes, for a given realization of the network

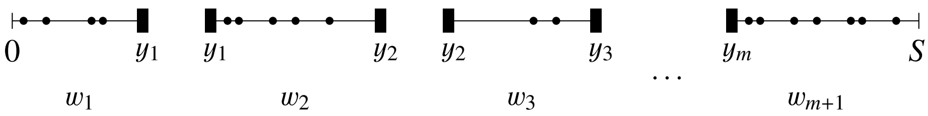

Figure 2.7: Decomposition of a network over $[0, S]$ into independent segments.

$$\Pr\{C\} = \Pr\{C_1\}\Pr\{C_2\}\cdots\Pr\{C_{m+1}\} \tag{2.6}$$

Conditioned on the number of nodes k_i in w_i, their distribution is uniform over w_i. This means that the position x_j of a given node in w_i has a distribution $\mathcal{U}(\|w_i\|)$. Where $\|w_i\|$ is the length of w_i. By the law of total probability, then (2.6) is given by

$$\begin{aligned}\Pr\{C\} = \sum_{k_1=0}^{n}\sum_{k_2=0}^{n-k_1}\cdots\sum_{k_m=0}^{n-k_1-\cdots-k_{m-1}}\binom{n}{k_1}\binom{n-k_1}{k_2}\cdots\binom{n-k_1-\cdots-k_{m-1}}{k_m}\\ \cdot\prod_{i=1}^{m+1}\Pr\{C_i \mid k_i \text{ in } w_i\}\Pr\{k_i \text{ in } w_i\}\end{aligned} \tag{2.7}$$

with

$$k_{m+1} = n - \sum_{i=0}^{m} k_i$$

and

$$\binom{n}{k}$$

representing the binomial coefficient, n choose k. Since the term

$$\Pr\{k_i \text{ in } w_i\} = \left(\frac{\|w_i\|}{\|\mathcal{T}\|}\right)^{k_i}$$

then to solve (2.7) we only require to calculate the terms $\Pr\{C_i \mid k_i \text{ in } w_i\}$.

Denote by $CB_{k,n,z}$ and by $CE_{k,n,z}$ the events of having border- and inner-connectivity respectively, using k out of n nodes in a segment z.

Let $C_{n,z}$ be the event that the network over segment length z with n nodes is connected. It was shown in [21, 35] that

$$\Pr\{C_{n,z}\} = \sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j \frac{(z-jr)^n}{z^n} \mathbf{1}_{\{z \geq jr\}} \tag{2.8}$$

where $\mathbf{1}_{\{\cdot\}}$ represents the indicator function

$$\mathbf{1}_{\{w \geq r\}} = \begin{cases} 1, & \text{if } w \geq r; \\ 0, & \text{otherwise.} \end{cases}$$

More generally, it was also shown in [35] that given a subsegment $[t, t+v] \subseteq [0, z]$ with n nodes in it, two of them fixed at the borders of the segment, the probability that the remaining $n-2$ nodes connect these border nodes is

$$\Pr\{CE_{n,v}\} = \sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j \frac{(v-jr)^{n-2}}{z^n} \mathbf{1}_{\{v \geq jr\}} \tag{2.9}$$

If we consider the event $CB_{k,n,z}$, then (2.8) can be extended to this more restrictive situation.

Theorem 2.2. *The probability of the event $CB_{k,n,z}$ for given $z = \|w_i\|$, $i \in \{1, m+1\}$ is*

$$\Pr\{CB_{k,n,z}\} = \frac{1}{\binom{n}{k}} \sum_{j=0}^{n} \binom{k}{j} (-1)^j \frac{(z-jr)^n}{z^n} \mathbf{1}_{\{z \geq jr\}} \tag{2.10}$$

Proof. Let $B = CB_{k,n,z}$; this event is equivalent to having $k+1$ nodes connected. Without loss of generality take $w = [0, z]$ and $y = 0$. Let u be the position of the rightmost node and note there are k possible options for the leftmost one. Using (2.9) we have

$$\begin{aligned} \Pr\{B\} &= \frac{k}{z^{n-k}} \int_0^z \Pr\{CE_{k+1,u}\} (z-u)^{n-k} \, \mathrm{d}u \\ &= \frac{k}{z^n} \int_0^z \sum_{j=0}^{n} \binom{k}{j} (-1)^j (u - jr)^{k-1} \mathbf{1}_{\{u \geq jr\}} (z-u)^{n-k} \, \mathrm{d}u \end{aligned} \tag{2.11}$$

Changing the order of the sum and integral and then by change of variables

$$t = \frac{u - jr}{z - jr}$$

it follows that

$$\Pr\{B\} = \frac{k}{z^n} \sum_{j=0}^{n} \binom{k}{j} (-1)^j (z - jr)^n \mathbf{1}_{\{z \geq jr\}} \int_0^1 t^{k-1} (1-t)^{n-k} \, \mathrm{d}t$$

The integral can be evaluated using the beta function B defined, for $x, y \in \mathbb{R}^+$, by

$$\mathrm{B}(x, y) = \int_0^1 (s)^{x-1} (1-s)^{y-1} \, \mathrm{d}s$$

that simplifies, for $a, b \in \mathbb{Z}^+$, to

$$\mathrm{B}(a, b) = \frac{(a-1)! \, (b-1)!}{(a+b-1)!}$$

Then the integral equals

$$\int_0^1 t^{k-1}(1-t)^{n-k}\,\mathrm{d}t = \frac{(k-1)!\,(n-k)!}{n!}$$

and the result follows. □

The event $CB_{k,n,z}$ occurs at segments w_i, $i \in \{1, m+1\}$, then it remains to provide a formula for $CE_{k,n,z}$, for w_i with $i \in \{2, 3, \ldots, m\}$. It suffices to have a formula for $CE_{n,n,z}$.

Theorem 2.3. *The probability of the event $CE_{n,n,z}$ for given $z = \|w_i\|$, $i \in \{2, 3, \ldots, m\}$ is*

$$\begin{aligned}\Pr\{CE_{n,n,z}\} = \frac{1}{z^n}\sum_{k=0}^{n}\sum_{j=0}^{k}\sum_{i=0}^{n-k}\binom{k}{j}\binom{n-k}{i}(-1)^{j+i}\,[z-(1+j+i)r]^n\,\mathbf{1}_{\{z\geq(1+j+i)r\}} \\ + \Pr\{CE_{n+2,z}\}\end{aligned} \tag{2.12}$$

Proof. Fix z, n. Let $E = CE_{n,n,z}$. Take A to be the event all nodes connect both sinks and B the event that each node reaches exactly one sink. Since the events are disjoint we have $\Pr\{E\} = \Pr\{A\} + \Pr\{B\}$. Where $\Pr\{A\}$ is given by (2.9) with $\Pr\{A\} = \Pr\{CE_{n+2,z}\}$. For B to occur there must be a segment of length greater than r without nodes. Thus we can consider the border connectivity of two segments, one with a sink at its left, the other with a sink at its right. Let B_1 be the event that k nodes are connected to the left sink and u is the location of the rightmost of these nodes—there are k possible combinations for this event. By (2.9) we have that $\Pr\{B_1\} = \Pr\{CE_{k+1,u}\}$, $u \in [0, z-r]$. Let event B_2 be that when there are $n-k$ nodes connected to the right sink within the subsegment $(u+r, z]$. Then we can use Theorem 2.2 to get $\Pr\{B_2\} = \Pr\{CB_{n-k,n-k,z-u-r}\}$. Therefore, conditioned on u, it follows

$$\Pr\{B \mid u\} = k \sum_{k=0}^{n} \binom{n}{k} \Pr\{B_1\} \Pr\{k+1 \text{ in } [0,u]\} \Pr\{B_2\} \Pr\{n-k \text{ in } (u+r, z]\}$$

Now since $u \in [0, z-r]$ we have

$$\Pr\{B\} = k \int_0^{z-r} \Pr\{B \mid u\} \, \mathrm{d}u \tag{2.13}$$

Using (2.9) and (2.10) and the fact that $k \sim \mathcal{U}[0,z]$ we have

$$\Pr\{B\} = \frac{k}{z^n} \int_0^{z-r} \sum_{k=0}^{n} \sum_{j=0}^{k} \sum_{i=0}^{n-k} \binom{n}{k} \binom{k}{j} \binom{n-k}{i} (-1)^{j+i} (u-jr)^{k-1} \cdot (z-u-r-ir)^{n-k} \mathbf{1}_{\{jr \le u \le z-r-ir\}} \, \mathrm{d}u$$

Changing the order of sums and integral, by Fubini's Theorem [14], and since $jr \le u \le z-r-ir$

$$\Pr\{B\} = \frac{k}{z^n} \sum_{k=0}^{n} \sum_{j=0}^{k} \sum_{i=0}^{n-k} \binom{n}{k} \binom{k}{j} \binom{n-k}{i} (-1)^{j+i} \int_{jr}^{z-r-ir} (u-jr)^{k-1} (z-u-r-ir)^{n-k} \, \mathrm{d}u$$

By the change of variable

$$s = \frac{u - jr}{z - (1+j+i)r}$$

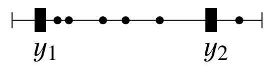

Figure 2.8: Representation of a network with two sinks.

we obtain

$$\Pr\{B\} = \frac{k}{z^n} \sum_{k=0}^{n} \sum_{j=0}^{k} \sum_{i=0}^{n-k} \binom{n}{k} \binom{k}{j} \binom{n-k}{i} (-1)^{j+i} [z - (1 + j + i)r]^n \int_0^1 (s)^{k-1} (1-s)^{n-k} \, \mathrm{d}s$$

where the integral is the beta function $\mathrm{B}(k, n-k+1)$ and evaluates to

$$\begin{aligned} B(k, n-k+1) &= \frac{\Gamma(k)\,\Gamma(n-k+1)}{\Gamma(n+1)} \\ &= \frac{(k-1)!\,(n-k)!}{n!} \\ &= \frac{1}{k\binom{n}{k}} \end{aligned}$$

The results of the theorem follows by combining the above formulas.

□

As an example of the use of (2.7), consider a network with two sinks at $y_1, y_2 \in [0, S]$, $S \in \mathbb{R}^+$, as illustrated in Fig. 2.8. Let C_{2S} be the event that the resulting network is connected, then we have

$$\begin{aligned} \Pr\{C_{2S}\} = \frac{1}{S^n} \sum_{k=0}^{n} \sum_{m=0}^{n-k} \binom{n}{k} \binom{n-k}{m} \Pr\left\{CB_{k,k,y_1}\right\} \Pr\left\{CE_{m,m,y_2-y_1}\right\} \\ \cdot \Pr\left\{CB_{n-k-m,n-k-m,S-y_2}\right\} y_1^k (y_2 - y_1)^m (S - y_2)^{n-k-m} \end{aligned} \tag{2.14}$$

Table 2.2: Parameters used in the simulation setup.

Parameter	Value
y	0.5
n	{9, 10}
$\mathcal{I}$	$[0, 1]$
(y_1, y_2)	{(0.25,0.85), (0.0,1.0)}
y_0	{0.50}
reps	10^5
r	(0,1)

2.4.3 Simulation Results

This Section presents simulation results that validate (2.7). To make the simulation concrete we consider the particular case given by (2.14) over $\mathcal{I} = [0, 1]$. The procedure for the simulations is by using the Monte Carlo method with 10^4 random replications for each set of parameters. Table 2.2 provides the simulation parameters.

The stochastic simulation method is as follows. For a given number of nodes n and a communication radius r, take a deployment over $\mathcal{I}$ and verify if the resulting network is connected. Repeat the procedure for the required number of replications for each r of interest. Then calculate $\Pr\{C_{2S}\}$ as the ratio of the number of connected network realizations and the number of replications for each corresponding radius r.

By using (2.14) we calculate the probability of connectivity for different r. Fig. 2.9 shows the probability of connectivity as a function of r with $n = 10$ for one sink location. Fig. 2.10 shows the probability of connectivity as a function of r with $n = 9$ for two sets of pairs of sink

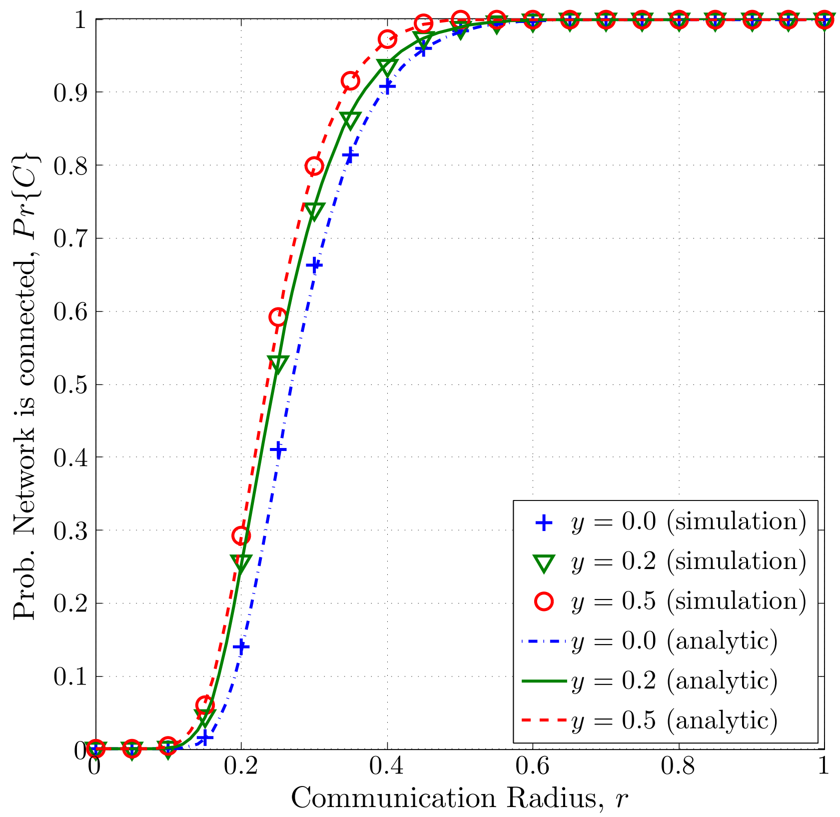

Figure 2.9: Probability of having a connected network with a given number of nodes and exactly one sink at location y.

locations. The first set of sinks are located at $y_1 = 0.25$ and $y_2 = 0.85$, while the second are at $y_1 = 0$ and $y_2 = 1$.

For comparison, and using (2.7), Fig. 2.10 also shows the probability of connectivity when there is one sink located at $y = 0.5$—the optimal location of a single-sink network—and $n = 10$—to compensate for the missing sink. Note the large increase in the probability—up to 53%—of having a connected network when comparing a deployment with two sinks versus one with just one sink.

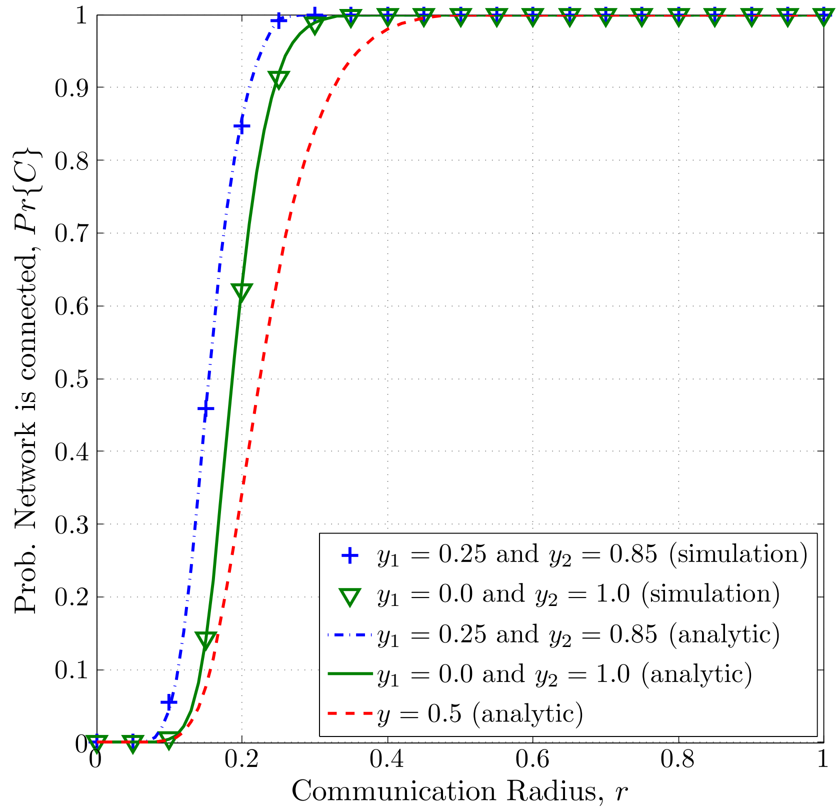

Figure 2.10: Probability of having a connected network with a given number of nodes and two sinks at locations y_1 and y_2.

2.5 Summary

This Chapter presented the connectivity analysis in one-dimensional wireless sensor networks in two settings: infrastructure-less and with infrastructure.

First, this Chapter analyzed the connectivity of WSNs with nodes having random communication radii. We focused on 1-D networks with finite number of nodes deployed uniformly over a line segment, each node having an independent and identically distributed random communication radius. Although it is a simple model, it can be used to account for more complex characteristics of the wireless channel. For example by modeling the randomness in the channel—as it was shown for the log-normal shadowing model when nodes have uniform communication radii. For this model, we presented a formula to calculate the probability of a source being able to send a message to a sink. Through simulations it was shown the validity of the analysis and the close proximity of our model to a WSN under log-normal shadowing, thus presenting an approximate but mathematically tractable scheme. We also noted a small increase in the probability of having a connected network when considering node platforms with random communication radii versus homogeneous radii.

The second main part of this Chapter presented an exact formula for the probability of having a connected 1-D WSN with infrastructure. Assuming the infrastructure to be a backbone of sinks, the network is connected if every node reaches at least one sink. This setting can be seen as a model of connectivity for WSNs having more than one connected component. We focused on 1-D networks with finite number of randomly deployed nodes having sinks at given arbitrary locations. We presented a formula for the probability of connectivity of such model. As a particular example of application of our contribution, a large increase in the probability of network connectivity was noted when comparing a deployment with two sinks versus a deployment with just one sink.

After considering the connectivity of WSNs in highly linear structures, modeled by 1-D networks, a natural extension of this problem is to analyze 2-D WSNs. That is the focus of the following Chapter.

CHAPTER 3

CONNECTIVITY ANALYSES IN 2-D WIRELESS SENSOR NETWORKS

3.1 Introduction

In the previous Chapter we considered connectivity in 1-D wireless sensor networks. In this Chapter we focus on a more general environment of a higher dimensionality, in particular in two-dimensional (2-D) wireless sensor networks. These 2-D WSNs have widely spread applications in sensing and monitoring fields. Two prominent examples of applications are outdoor fields—such as wild animal habitats—and physical structures—such as buildings and production floors.

We focus on the problem of network connectivity when having random deployment of nodes over the field of interest. Specifically, consider a certain number—random or deterministic—of nodes randomly deployed in a 2-D finite area. For the analysis of such network we will consider both a deterministic and a random communication channel—log-normal shadowing effect—as well as a network with infrastructure. Note that a 2-D setup is a complex environment and its exact analysis is an open problem, here we make some simplifying assumption to obtain approximate formulas for the probability of network connectivity. However, as we show below through stochastic simulations, our approximations are useful, in the sense that their predicted outcomes are very close to the results provided by the network simulations.

To address the problem of network connectivity, in this Chapter we use the theory of stochastic geometry along with principles of random geometric graphs. Our key results are the following. First we provide formulas for the probability of connectivity in 2-D WSNs under a simple communication link model—Boolean—and when the nodes are distributed randomly over the field of interest. Then we provide formulas for the required density of nodes to be deployed in order to keep a WSN connected with a given probability when such network is analyzed under the log-normal shadowing model for the communication channel. Lastly, we use and analyze the concept of partial connectivity in a random WSN with infrastructure. Partial connectivity allows the existence of few isolated node platforms and thus can improve other network metrics. This concept of partial connectivity, as discussed below, can be coupled with the use of infrastructure in a given field of interest.

This Chapter is organized as follows. We divide the analysis of the network connectivity by the characteristics of the communication channel and by the level of connectivity. The Section 3.2 presents the related work in the literature. In §3.3 we cover the Boolean model, while the subsequent Section 3.4 considers the communication channel under slow fading, specifically using the log-normal shadowing model. The last Section 3.5 introduces and analyzes the concept of partial connectivity in a random network with infrastructure.

3.2 Related Work

There are multiple results and various approaches that address the topic of connectivity of wireless networks. The following works, representative research related to this Chapter, have addressed full and partial connectivity of random wireless networks in 2-D deployments.

The work in [38] presents a study on 2-D dense networks with n uniformly distributed nodes. Using percolation theory, the authors consider the communication radius required for asymptotic connectivity.

In [12] there is an analysis of connectivity under log-normal shadowing model where nodes are deployed over an infinite plane according to a Poisson process. The authors provide bounds on the probability of connectivity using random geometric graphs theory. The work in [23] considers connectivity issues under interference from neighboring nodes, but does not take into account fading in the communication channel.

Regarding hybrid network deployments and partial connectivity, the more relevant research to this dissertation is the following. As the opposite to 2-D WSNs, the asymptotic results in 1-D wireless networks show there is no *Percolation*: it is not possible to have an infinite size connected component, not even with base stations on a lattice [25]. That work analyzes wireless networks with constant density through a Poisson Boolean model and it estimates the probability that an arbitrary node belongs to an infinite size connected component.

In [11] there is an analysis of connectivity where nodes are deployed over an infinite plane according to a Poisson point process. The work provides bounds on the probability of connectivity considering a finite sub-area and using approximations based on geometric random graphs. It is also possible to use Percolation theory to analyze infinite 2-D networks and obtain the critical node density that supports an infinite connected component almost surely [25].

There are few connectivity results in the literature for hybrid 2-D networks. An example is also [25] that uses Percolation theory, as mentioned above. The work in [32] analyzes the connectivity within a finite area of an infinite plane. Both nodes and sinks are deployed following independent Poisson point processes and data communication links are restricted to single-hop. The model also takes into consideration the interference among nodes of the wireless network. By only considering single-hop communication links, [27] uses a statistical approach to approximate the probability of connectivity through the probability of a node not being isolated. The authors consider a network deployment with multiple sinks over a finite region.

The research in [24] provides insights about the partial connectivity in 2-D wireless networks where, asymptotically, all disconnected nodes are isolated nodes, this means that the most isolated node determines the critical range for full connectivity.

3.3 WSNs with Deterministic Communication Channel

In this Section the connectivity analysis is performed under the Boolean model while in the following Section it will be under the log-normal shadowing communication channel model. We provide formulas for the probability of network connectivity as a function of the density of the node platforms deployment.

Let us consider a 2-D WSN consisting of randomly deployed nodes over a field of interest with finite area. See Fig. 3.1 for an illustration of a network realization. The method to address the connectivity problem is based on the theory of point processes as well as some results from random geometric graphs.

An example of application of the formulas for the probability of connectivity in 2-D WSNs is the framework of monitoring physical structures with WSNs. This is due to the idea that node platforms will be embedded into structures—like homes, office buildings, and power grids—following a random deployment [26].

3.3.1 Model and Problem Formulation

Denote the field of interest by an open rectangle $S = (-w, w) \times (-d, d)$ in $\mathbb{R}^2$. Consider a static WSN with a random deployment of nodes over S following a Poisson process [45] with intensity $0 < \lambda < \infty$ with units $[\text{area}]^{-1}$, denote it by Π_λ. Note that this is a particular case of a Point Process [19]. To avoid border effects, consider S to form a torus and

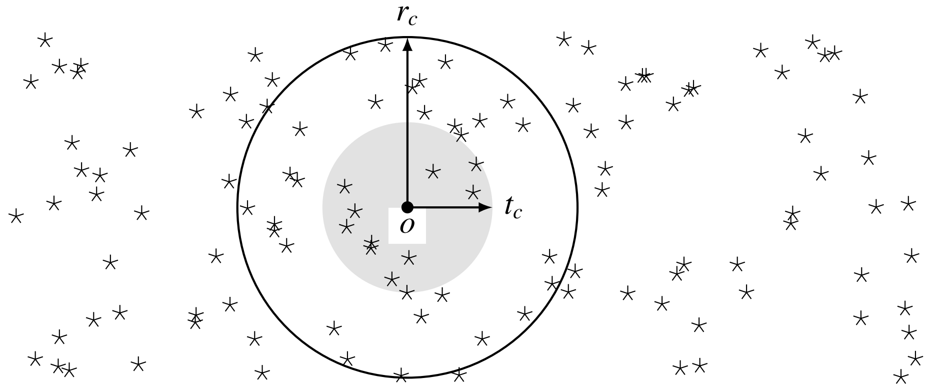

Figure 3.1: Example of a realization of a network. Node o has communication radius r_c, when $\sigma = 0$. The radius where there is perfect correlation is t_c.

measure the distance between a pair of nodes accordingly. Let $N(B)$ be a random variable that represents the number of nodes in the set $B \subseteq S$. Then $N(B)$ has a Poisson distribution with parameter λ, that is

$$\Pr\{N(B) = n\} = \frac{\lambda^n}{n!} \mathrm{e}^{-\lambda} \tag{3.1}$$

Denote by $\boldsymbol{x}_i \in S$ the position of node i. Without loss of generality (wlog), assume there is a node o at the origin $(0, 0)$. Indeed, when considering the field of interest as a torus, we can use the Slivnyak's theorem, and o can be taken either arbitrarily in S or at an arbitrary point in Π_λ [74]. This means that the Poisson process seen from the position of a randomly chosen point has the same distribution as the original process plus that point.

Let C be the event that the network is connected, then the problem statement is the following

Given the above conditions, what is the density λ of nodes required for $\Pr\{C\} = p$, $0 < p < 1$?

First we solve for a general setting, then specialize the results for the two models of interest for the communication channel—Boolean links and log-normal shadowing.

3.3.2 Analysis

Consider the deployment of node platforms according to the process Π_λ. Let NI be the event that no node in S is isolated. By random geometric graph theory [61, Chapter 13], given $N(S) = n$, it follows that

$$\Pr\{NI\} \to \Pr\{C\} \text{ as } n \to \infty$$

Then when λ is large, $N(S)$ will be large and $\Pr\{C\} \approx \Pr\{NI\}$. Thus we can analyze $\Pr\{NI\}$. Take $N(S) = n$, $n \in \mathbb{N}$. Label each node by $i \in \{1, 2, \ldots, n\}$. Let I_i be the event that node i is isolated and NI_i that it is not, then

$$\begin{aligned} \Pr\{NI \mid N(S) = n\} &\overset{(a)}{=} \Pr\{NI_i, \forall i \mid N(S) = n\} \\ &\overset{(b)}{\approx} (\Pr\{NI_1\})^n \\ &= (1 - \Pr\{I_1\})^n \end{aligned} \tag{3.2}$$

where (a) follows by definition, while (b) follows since when λ is large the minimum largest communication link between two nodes required for having a connected network will be relatively small [61, Chapter 2]. Therefore the probability that two nodes are isolated is almost an independent event. Using (3.2) and by the law of total probability

$$\begin{aligned}\Pr\{NI\} &= \sum_{n=0}^{\infty} \Pr\{NI \mid N(S) = n\} \Pr\{N(S) = n\} \\ &\approx \sum_{n=0}^{\infty} (1 - \Pr\{I_1\})^n \frac{(\lambda S)^n}{n!} e^{-\lambda S} \\ &= e^{-\lambda S} e^{\lambda S(1 - \Pr\{I_1\})} \\ &= e^{-\lambda S \Pr\{I_1\}}\end{aligned} \tag{3.3}$$

Thus since $\Pr\{C\} \approx \Pr\{NI\}$, using (3.3) we have

$$\Pr\{C\} \approx e^{-\lambda S \Pr\{I_1\}} \tag{3.4}$$

Hence to obtain the approximation of $\Pr\{C\}$ we need $\Pr\{I_1\}$ or, by the properties of Poisson point processes, equivalently $\Pr\{I_o\}$. This probability is a function of the existence of a communication link between nodes. Denote by $\mathbb{E}(\|G\|) < \infty$ the expected *content* of an arbitrary random shape $G \in S$. As shown in [40, Chapter 3] it follows

$$\Pr\{I_o\} = e^{-\lambda \mathbb{E}(\|G\|)} \tag{3.5}$$

Note that I_o depends only on the expected content of the shape of interest, thus by symmetry and homogeneity of the Poisson process the same holds for any point i in S. We specialize (3.5) to a communication channel with deterministic power path-loss in this Section and with log-normal shadowing in the following Section.

Boolean Communication Links

Assume there is a deterministic path-loss in the communication channel and each node has a homogeneous transmission power such that its communication radius r_c is r. Let $\mathrm{b}(\boldsymbol{x}, r)$ be a disc with center $\boldsymbol{x} \in S$ and radius r. Node o transmits and node i at $\boldsymbol{x}_i$ receives if and only if their distance is less than r, that is $\|\boldsymbol{x}_i\| \leq r$. Then for o to be isolated, there should be no other node within $\mathrm{b}(\boldsymbol{x}_o, r)$, that is

$$\begin{aligned} \Pr\{\, I_o \,\} &= \Pr\{\, N(\,\mathrm{b}(\boldsymbol{x}_o, r)\,) = 0\,\} \\ &= \mathrm{e}^{-\lambda\pi r^2} \end{aligned} \tag{3.6}$$

By using (3.6) in (3.4) the probability of connectivity for the Boolean communication links model is approximated by

$$\Pr\{\, C_B \,\} \approx \exp\left\{-\lambda S\, \mathrm{e}^{-\lambda\pi r^2}\right\} \tag{3.7}$$

3.3.3 Simulation Results

This Section presents simulation results that validate the developed analytic formula for the network and channel model in the previous Section. The procedure for the simulations is by using the Monte Carlo

Table 3.1: Parameters used in the simulation setup.

Parameter	**Value**
reps	10^4
λ	$(0, 1600)$
S	$[0,1] \times [0,1]$
r_c	$\{0.6,\ 0.7,\ 0.8\}$

method with 10^4 random replications for each set of parameters. Table 3.1 provides the simulation parameters.

The simulation technique is as follows. Take a deployment over S, with a density of nodes λ and verify if the resulting network is connected. Repeat the procedure for the required number of replications for each λ of interest. Then to calculate the $\Pr\{C\}$ take the ratio of the total number of connected network realizations and the total number of replications for each corresponding λ. For clarity purposes the confidence intervals are not depicted in the resulting graphs. However we remark that they were small—the maximum deviation width was less than 0.7% of the estimated average value of $\Pr\{C\}$.

Using results from §3.3.2 we calculate $\Pr\{C\}$ for different λ. Fig. 3.2 shows the simulation results. Solid lines represent the probability of no node being isolated, as provided by (3.7). Markers represent the probability of having a connected network via simulation. Notice that for large $\Pr\{C\}$—the values of practical interest, the approximation between no isolated node and having a connected network is very accurate.

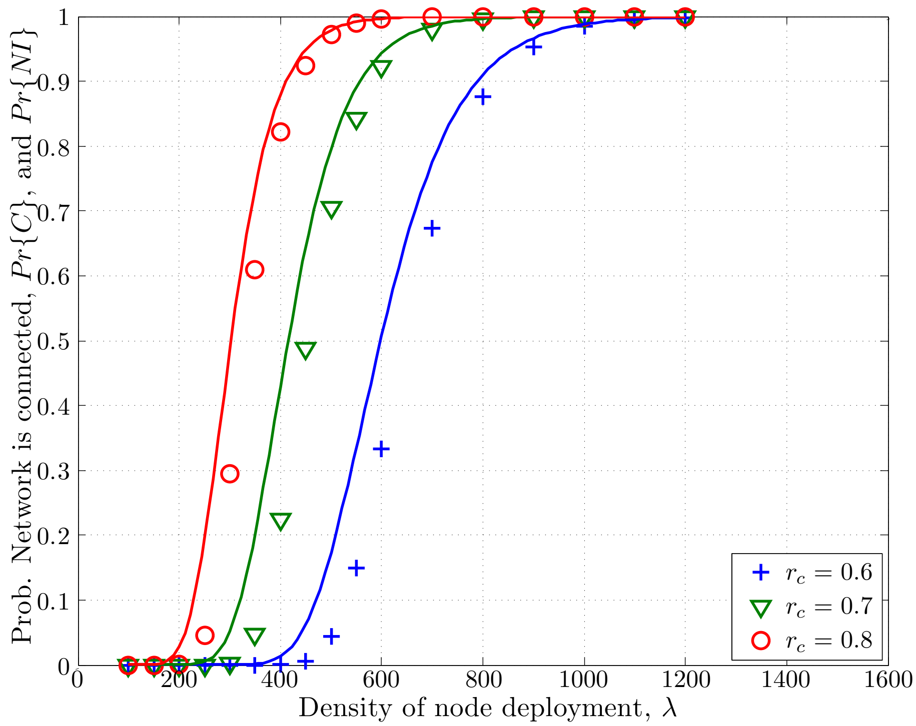

Figure 3.2: Probability of having a connected network and having no isolated nodes as a function of the density of the deployed nodes for different r_c.

3.4 WSNs with Random Communication Channel

In this Section we consider a more complex environment than that of the Boolean communication links. In particular, we analyze a model that takes into consideration the communication channel randomness. We provide formulas for the probability of network connectivity as a function of the density of the node platforms deployment.

This type of analysis under communication channel with randomness is essential for real world environments. A general model of the noise present in electronic equipment is the Additive White Gaussian noise (AWGN). For this reason we consider the log-normal shadowing model, since it is a widely accepted statistical model used to analyze such randomness in the system's channel [20, 75]. The shadowing represents the slow fading of the radio signals and, with the appropriate parameters, it is valid for either indoor or outdoor environments [63].

As in the previous Section, let us consider a 2-D WSN consisting of randomly deployed nodes over a field of interest with finite area. Refer back to Fig. 3.1 for an illustration of such network realization.

3.4.1 Model and Problem Formulation

Consider the same assumptions as in §3.3.1. We mention them briefly. S denotes the field of interest, an open rectangle. Nodes are deployed over S following a Poisson process Π_λ with intensity $0 < \lambda < \infty$. Let us use a torus distance for the communication links. $N(B)$ is the number of nodes in set $B \subseteq S$.

Let C be the event that the network is connected, then the problem statement is:

> *Given the above conditions and assuming shadowing in the communication channel, what is the density λ of nodes required for* $\Pr\{C\} = p$, $0 < p < 1$?

3.4.2 Analysis

Recall, from (3.4), that in order to obtain the approximation of $\Pr\{C\}$ we need to calculate $\Pr\{I_o\}$. And this probability is a function of the existence of a communication link between nodes. Next, we use (3.5) in a communication channel modeled by log-normal shadowing.

Shadowing in the Communication Channel

In this model of a wireless channel besides the deterministic signal attenuation, denoted by η, the path-loss exponent, there is a random attenuation due to obstructions in the communication links. That randomness is modeled by a random variable. The formula for the total path-loss L in decibels [dB] for a given distance l in meters [m] between transmitter and receiver is

$$L(d,\sigma) = \bar{L}(d_0) + 10\eta \log\left(\frac{d}{d_0}\right) + X_\sigma \qquad \text{[dB]} \tag{3.8}$$

where $l_0 \ll l$ [m] is the reference distance where the reference power $\bar{L}(l_0)$ [dB] is measured. X_σ is a normally distributed random variable $\mathcal{N}(0,\sigma^2)$ [dB]. Hence L is also a random variable. In practical environments the values of σ fall within $(1,10)$, while those of η range within $(2,6)$ [63].

When considering the shadowing model, a node located at distance l from the transmitter will be able to correctly receive the information if the received power of the signal is above certain threshold or, equivalently, if the total path-loss is less than a particular threshold L_{th}, that is there is a communication link if and only if $L(l,\sigma) \leq L_{th}$. Denote by $e(l,\sigma)$ the event that there is a link between two nodes at a distance l when using the parameter σ for the shadowing model. Without loss of generality let $\bar{L}(l_0) = 0$ [dB] and $l_0 = 1$ [m], then

$$\begin{aligned}
\Pr\{\, e(l,\sigma)\,\} &= \Pr\{\, L(l,\sigma) \leq L_{th}\,\} \\
&= \Pr\{\, 10\eta \log l + X_\sigma \leq L_{th}\,\} \\
&= \Pr\{\, X_\sigma \leq L_{th} - 10\eta \log l\,\} \\
&= \Phi\left(\frac{L_{th} - 10\eta \log l}{\sigma}\right)
\end{aligned}$$

Where $\Phi(\cdot)$ is the cumulative function of the standard normal distribution. Denote by $e(\sigma)$ the content of the random shape formed by the possible communication links around o. We need to obtain $e(\sigma)$ in order to use (3.5). Note that

$$\begin{aligned}
\mathbb{E}\{\, e(\sigma)\,\} &= \lambda \int_0^{2\pi}\int_0^{\infty} \Pr\{\, e(l,\sigma)\,\}\; l \,\mathrm{d}l \,\mathrm{d}\phi \\
&= 2\pi\lambda \int_0^{\infty} \Phi\left(\frac{L_{th} - 10\eta \log l}{\sigma}\right) l \,\mathrm{d}l
\end{aligned} \tag{3.9}$$

By using (3.9) in (3.4), the probability of connectivity for the log-normal shadowing communication channel model is approximated by

$$\Pr\{\, C_S\,\} \approx \exp\left\{-\lambda S \,\mathrm{e}^{-2\pi\lambda \int_0^{\infty} \Phi\left(\frac{L_{th}-10\eta \log l}{\sigma}\right) l \,\mathrm{d}l}\right\} \tag{3.10}$$

Table 3.2: Parameters used in the simulation setup.

Parameter	**Value**
reps	10^4
λ	(0,1600)
σ	{4, 8}
η	3
S	$[0,1] \times [0,1]$
r_c	{0.7, 0.8}

3.4.3 Simulation Results

This Section presents simulation results that validate the developed analytic formula for the network and channel model in the previous Section. For the parameters of shadowing take $\sigma = 4$ and 8 and $\eta = 3$—typical values in real settings [63]. To make a fair comparison with the simulation results for the Boolean model in §3.3.3 define L_{th} to be such that $r_c = 0.7$ and 0.8 when $\sigma = 0$. Take $\bar{L} = 0$ and $d_0 = 1$. The procedure for the simulations is by using the Monte Carlo method—see Section 3.3.3 for a description of the mechanism—with 10^4 random replications for each set of parameters. Table 3.2 provides the simulation parameters.

Using results from §3.4.2 we calculate $\Pr\{C\}$ for different λ. Fig. 3.3 shows the simulation results. Solid lines represent the probability of no node being isolated, as provided by (3.10). Markers represent the probability of having a connected network via simulation for different values of σ and r_c. For comparison purposes, the broken lines with no markers represent the probability of no node being isolated when $\sigma = 0$, that is when the communication link is Boolean. Notice that for

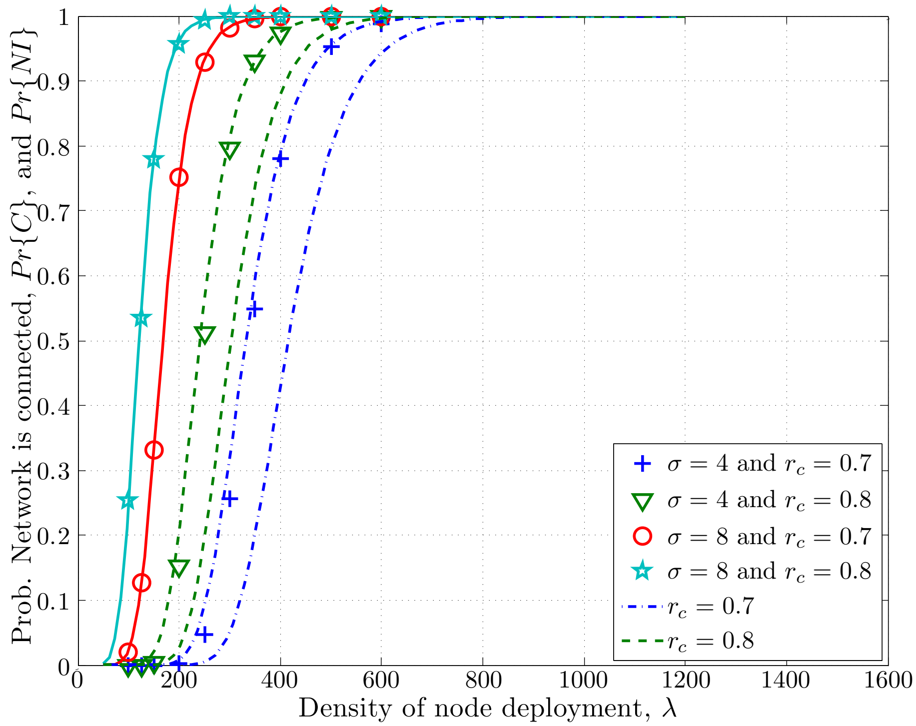

Figure 3.3: Probability of having a connected network and having no isolated nodes with respect to the density of deployed nodes under different r_c and σ.

large $\Pr\{C\}$, the approximation between no isolated node and having a connected network is very accurate. Also note that when σ increases, the approximation is better. Large values of σ increase connectivity, in agreement to previous results in two-dimensional networks [12].

3.5 Partial Connectivity of Hybrid WSNs

As we mentioned, connectivity is a fundamental prerequisite for wireless sensor networks functionality—sensors must be able to report their measurements or observations if they are to serve their intended purpose. There are several results in the literature on wireless networks that treat connectivity [7, 8, 12, 35, 38, 57]. Though hybrid networks can be very useful [70] in particular applications, only a few of these works consider sinks as a part of the network. See the work on hybrid networks [5, 25, 27, 31].

This Section considers a finite two-dimensional (2-D) wireless network consisting of finite number of randomly deployed nodes and base stations at fixed locations around the border of the network. It analyzes its probability of connectivity, that is the ability of each node to convey a message to at least one of the few sinks. See Fig. 3.4 for an illustration. In addition, this Section considers partial connectivity, where a network can have isolated nodes. Isolated nodes would be acceptable in a large network of redundant sensors. The networks are analyzed under the Boolean communication link model and nodes and sinks are assumed to have homogeneous communication radii.

An example of application of the results from this Section is in the framework of monitoring physical structures with randomly deployed WSNs. In this setting an architecture that has several sinks and allows for partial connectivity may prove advantageous. Note that a WSN with more than one sink can be more practical for monitoring large areas [31]; also, by having various points of fusion, the network can be more reliable, and complex, than with one sink. Allowing for partial connectivity improves WSNs scalability and throughput [24] and nodes require less transmission power—thus having smaller size and/or longer network lifetime.

The results presented in this Section show the benefits of using in-

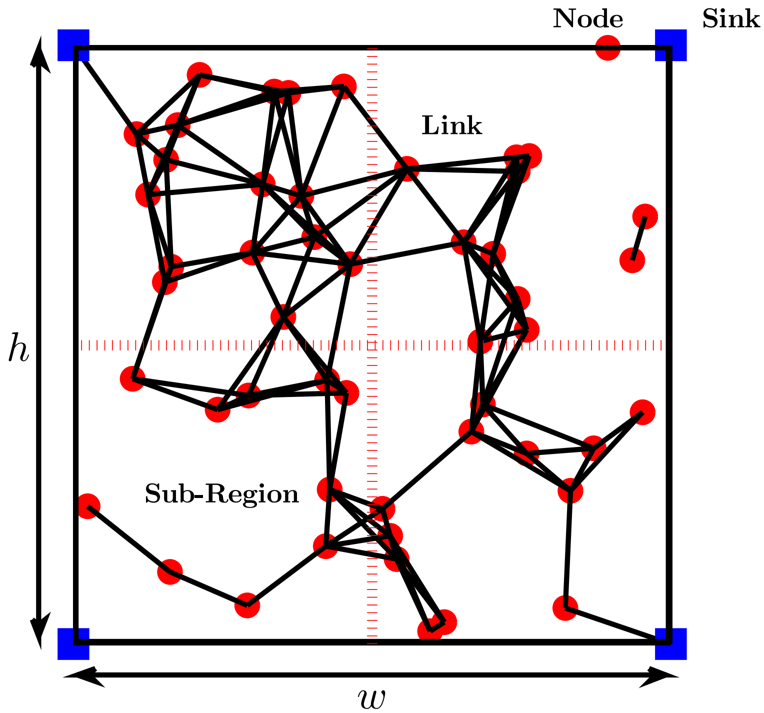

Figure 3.4: Example showing a realization of a network and its partition in four equal-area sub-regions, each one including a sink.

frastructure and provide an approximate formula for the connectivity of such network when there are few sinks. Also we show how partial connectivity improves some network metrics at the cost of having a small number of isolated nodes. The contributions of this section include an insight on the fraction of nodes that reach a sink as well as a simple approximation of the probability of connectivity in WSNs with infrastructure.

As mentioned in §3.2 all the previous results from the literature consider full connectivity—any node is able to communicate to any other node, a base station, or belongs to an infinite component. Here we will consider partial connectivity. The partial connectivity is application dependent, for example in large networks of redundant sensors, but there are benefits in allowing it: improvements in network scalability, throughput, and resource use. When defining connectivity in the information theoretical sense if a network scales, then it is necessary to reduce the communication throughput to achieve full connectivity [24].

3.5.1 Model and Problem Formulation

The general model for this Section is as follows. Consider a multi-hop 2-D finite hybrid WSN. Let us focus on few sinks located at the border of the field of interest. This location has practical relevance; for example it applies when a network monitors physical structures—like a bridge—where sinks are placed at the perimeter of the structure while the nodes are randomly deployed inside it.

Let the rectangle $\mathcal{F} \in \mathbb{R}^2$ represents the field of interest, with area $||\mathcal{F}|| = F = w \cdot h$. Consider a uniform deployment of n nodes over $\mathcal{F}$ and assume there are m sinks symmetrically located at the border of $\mathcal{F}$, such that its equipartition is possible. Fig. 3.4 illustrates $w = h$ and $m = 4$. Further let links be Boolean and both sinks and nodes have same communication radius r_c.

The problem statement is the following

> *Let $\mathcal{F}$ be the field of interest, n the number of nodes, and m the number of sinks, with characteristics defined above. Denote by FC_n^m the event that this network is fully connected—each node*

can reach at least one sink. Then what is the probability of the event FC_n^m?

3.5.2 Full Connectivity Analysis with no Sinks

Let us start with the analysis of the full connectivity of a WSN with no sinks and then get extensions of the results. We want the probability of connectivity of n nodes and m sinks in $\mathcal{F}$. There is no an exact connectivity formula in 2-D networks, but it can be approximated fairly close, as described below.

The probability of connectivity is close, and smaller, to the probability that there are no isolated nodes in $\mathcal{F}$. In turn, this probability of no isolated nodes can be approximated by knowing the probability that a single node is isolated. Below we develop and justify these ideas.

Denote by FC_n the event that a network with n nodes and no sinks is fully connected. Let NI be the event that no node in $\mathcal{F}$ is isolated. Label each node by $i \in \{1, 2, \ldots, n\}$. Let I_i be the event that node i is isolated and NI_i that it is not, then

$$\begin{aligned} \Pr\{NI\} &= \Pr\{NI_i,\ \forall i\} \\ &\overset{(a)}{\approx} (\Pr\{NI_k\})^n \\ &= (1 - \Pr\{I_k\})^n \end{aligned} \tag{3.11}$$

where (a) follows since for large n the minimum largest link between two nodes required for having a connected network will be relatively small [61, Chapter 2]. Then the probability that two nodes are isolated is almost an independent event. Node k in (3.11) represents a *typical*—random—node, as defined in [74]. Now, since $\Pr\{I_k\}$ decreases monotonically when n increases [61], if $\Pr\{I_k\}$ is small and n large, then we can approximate (3.11) using the definition of the exponential function

$$\mathrm{e}^x = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n$$

and a change of variables. Thus

$$\Pr\{NI\} \approx \mathrm{e}^{-n\Pr\{I_k\}} \tag{3.12}$$

To get $\Pr\{I_k\}$ considering border effects in $\mathcal{F}$ we use a similar procedure as [11]. Given k's location $z = (x, y) \in \mathcal{F}$

$$\Pr\{I_k\} = \int_{\mathcal{F}} \Pr\{I_k \mid z\} f_Z(z)\,\mathrm{d}z \tag{3.13}$$

where f_Z is the generalized probability density function of the uniform distribution

$$f_Z(z) = \begin{cases} \frac{1}{F}, & \text{if } z \in \mathcal{F}; \\ 0, & \text{otherwise.} \end{cases}$$

and $F = \|\mathcal{F}\|$ is the area of $\mathcal{F}$.

Since $\Pr\{I_k\}$ is small and n large, we can approximate a single node being isolated by the Poisson distribution [74] as

$$\Pr\{\, I_k \mid z \,\} = \mathrm{e}^{-\lambda(z)} \tag{3.14}$$

here $\lambda(z)$ is the expected number of nodes in $\mathcal{F}$ within r_c of z

$$\lambda(z) = \int_{\mathcal{S}} n\, f_Z(z)\, \mathrm{d}z \tag{3.15}$$

where $\mathcal{S} = \{\, \mathrm{b}(z\,, r_c) \cap \mathcal{F} \,\}$ and $\mathrm{b}(z\,, r_c)$ is a disc with center $z \in \mathcal{F}$ and radius r_c. Thus by knowing $\lambda(z)$ we can obtain $\Pr\{\, NI \,\}$. From (3.13) and (3.14)

$$\Pr\{\, I_k \,\} = \int_{\mathcal{F}} \mathrm{e}^{-\lambda(z)} f_Z(z)\, \mathrm{d}z \tag{3.16}$$

From (3.12) and (3.16), and since $\Pr\{\, FC_n \,\} \approx \Pr\{\, NI \,\}$ for large n [61], the result follows

$$\Pr\{\, FC_n \,\} \approx \Pr\{\, NI \,\} \approx \exp\left(-n \int_{\mathcal{F}} \mathrm{e}^{-\lambda(z)} f_Z(z)\, \mathrm{d}z \right) \tag{3.17}$$

As an example of obtaining $\lambda(z)$, take n nodes, $w = h = 2l$, then $F = 4l^2$—other rectangular regions are treated similarly. Call *border* nodes the ones located within r_c of any border in $\mathcal{F}$, and *center* nodes the rest. Let A_c, $A_s(x)$, and $A_w(z)$ represent the area of a circle, a circular segment (the shadowed part of the circle in Fig. 3.5), and the intersection of two such segments, respectively. By symmetry, to get $\lambda(z)$ it suffices to consider the four regions indicated in Fig. 3.5. Region I contains the center nodes

$$\lambda_I(z) = \frac{nA_c}{F}$$

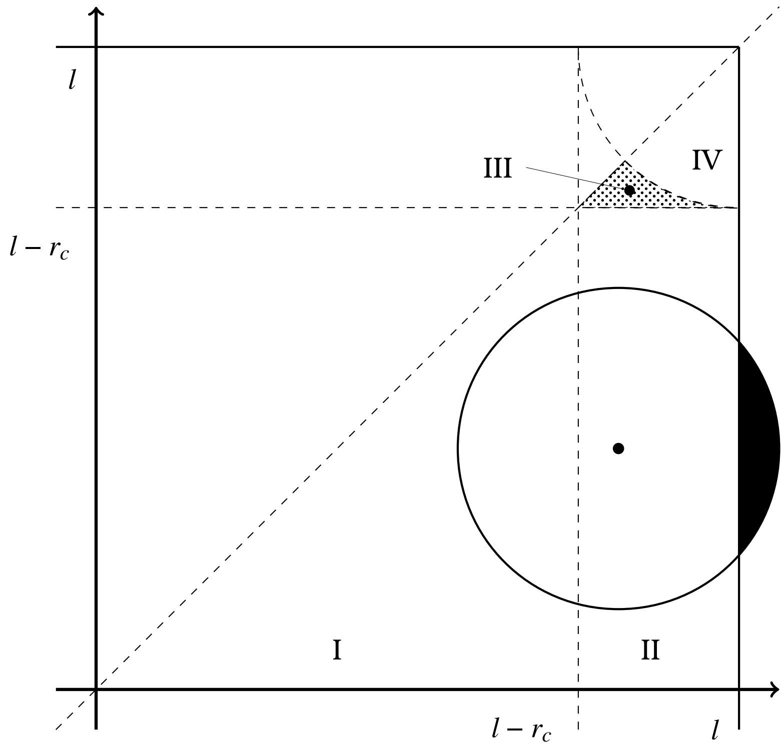

Figure 3.5: One quadrant of a square field of interest. The different regions used for the estimation of a node being isolated are shown.

For border nodes—regions II, III, IV—we have

$$\lambda_{II}(z) = \frac{n\,(A_c - A_s(x))}{F}$$
$$\lambda_{III}(z) = \frac{n\,(A_c - A_s(x) - A_s(y))}{F}$$
$$\lambda_{IV}(z) = \frac{n\,(A_c - A_s(x) - A_s(y) + A_w(z))}{F}$$

where $A_c = \pi r_c^2$ and, by setting $d(x) = \sqrt{(r_c^2 - (l-x)^2}$ and $s(x,y) = (\,d(x) - (l - y)\,)$, for position $z = (x, y)$,

$$A_s(x) = r_c^2 \cos^{-1}\left(\frac{l-x}{r_c}\right) - (l-x)\,d(x)$$
$$A_w(z) = \frac{s(x,y)\,s(y,x)}{2} + A_s\left(l - \frac{1}{2}\sqrt{4r_c^2 - s(x,y)^2 - s(y,x)^2}\right)$$

3.5.3 Full Network Connectivity with Sinks

Let us relate FC_n^m and FC_n. Clearly $\Pr\{FC_n^m\} \geq \Pr\{FC_n\}$. Since nodes are deployed uniformly over $\mathcal{F}$, the expected number of nodes in a given region $R \in \mathcal{F}$ is directly proportional of the area of the region $\|R\|$. By using this property of the uniform distribution, we define subregions of connectivity, each of equal area and with a sink, and from them we obtain $\Pr\{FC_n^m\}$. This is an approximation because connected components within a region may extend away of the region boundaries, as in Fig. 3.4. Moreover, the regions are not completely independent of each other, so the approximation is a lower bound that is mathematically tractable.

Next we justify our approximation. Label each sink by $j \in \{1, 2, \ldots, m\}$. Partition $\mathcal{F}$ in m disjoint regions of equal area A_j, such that $\sum_j A_j = F$ (the area of $\mathcal{F}$) and each region contains a sink j. Note that $\Pr\{FC_n\}$ in (3.17) is a function of the area F through the weighted integral. Thus we can express (3.17) as

$$\Pr\{FC_n\} = \exp\left(-\frac{n}{F}\,g(F)\right), \quad g(F) = \int_{\mathcal{F}} \mathrm{e}^{-\lambda(z)}\,\mathrm{d}z \tag{3.18}$$

Denote by $FC(j)$ the event that the nodes in A_j are connected to sink j. Then by the definition of the general uniform distribution (proportional to area) and linearity of integrals

$$\begin{aligned}\Pr\{FC(j)\} &\equiv \exp\left(-\frac{n}{F}\,g\left(\frac{F}{m}\right)\right) \\ &\equiv \exp\left(-\frac{n}{mF}\,g(F)\right)\end{aligned} \tag{3.19}$$

Hence, by the approximation of independence of regions, the probability that all regions are connected is

$$\begin{aligned} \Pr\{FC(j),\ \forall j\} &= (\Pr\{FC(1)\})^m \\ &= \Pr\{FC_n\} \end{aligned} \tag{3.20}$$

Hence by (3.20) it is possible to consider m regions, each with a sink, and use (3.17) to obtain an approximation of $\Pr\{FC_n^m\}$.

3.5.4 Partial Network Connectivity with Sinks

This Section describes the principles of partial connectivity of a hybrid wireless network under the model of §3.5.1. A simulation study approach below, in §3.5.5, presents the likelihood of partial connectivity.

Close inspection of the literature on the requirements of the pair (n, r_c)—number of nodes and their communication radius respectively—in a typical setting of random wireless sensor networks is insightful. It is observed that at least one of the quantities in such pair is quite large compared to the size of the field of interest and a structured node deployment. The reason for these large numbers is the randomness in the deployment of nodes and the strict requirement of having a fully connected network with high probability.

One problem of having large n and/or r_c in a relatively small field is the resulting large average node degree [24]. And, as a consequence, the interference between nodes is substantial and their communication rates are small—since there are more collisions, retransmissions, or delays. A major problem when having large r_c is the correspondent large power used by a node platform when communicating.

As noted, partial connectivity may be feasible in certain applications, such as large networks of redundant sensors. It is in these type of settings where a wireless sensor network becomes fully connected when the last isolated node becomes connected [61]. This suggests that the last isolated node has a very strong influence in the estimations on the total number of nodes to be randomly deployed to achieve connectivity in the first place.

In the partial connectivity scenario it is acceptable to have some isolated nodes in order to improve other metrics of a network, like interference due to node degree and energy savings by reducing the required minimum communication radius to have a functional network. Note that by adjusting n and/or r_c, the ratio of isolated nodes to the total number of nodes in a deployment can be as small as desired, up to the point where there is a full network connectivity with high probability. We consider network metrics like fraction of nodes reaching a sink and average node degree. As we show below, in some situations the improvement of the network metrics can be large.

Let us define a network as *well* or *partially connected* if only a very small fraction of the total number of nodes is isolated [24]. For the setting when there are multiple base stations, the likelihood of connectivity is the fraction of nodes that reach a sink, that is the probability

Table 3.3: Parameters used in the simulation setup.

Parameter	**Value**
reps	10^4
$\mathcal{F}$	$[0,1] \times [0,1]$
m	4
small n	
n	$\{20, 60, 100\}$
r_c	$(0.0, 0.6)$
large n	
n	$\{500, 1000\}$
r_c	$(0.0, 0.2)$

that there exists a path from *most* of the connected components to at least one of the multiple sinks. Fig. 3.4 shows a partially connected network. Practical deployments could require, for example, that at least 95% of the node platforms are connected or able to communicate to at least one base station.

3.5.5 Simulation Results

This Section presents simulation results that validate the formulas developed in the previous Sections. It also exhibits a comparison of the probability of connectivity between a WSN with sinks versus one without them, as well as a contrast between partial and full connectivity in a network with sinks. In particular, the number of sinks is $m = 4$, one at each corner of a square $\mathcal{F} = (0,1)^2$. We note that similar results hold with different symmetric location of the sinks along the border of $\mathcal{F}$. The simulations use the Monte Carlo method with 10^4 random replications for each set of parameters. Table 3.3 provides the simulation parameters.

The simulation mechanism is as follows. Take a uniform deployment of n nodes and verify if the resulting network is fully (or partially) connected. Repeat the procedure for the required number of replications for each n and r_c of interest. To calculate the probability of connectivity for a given pair (n, r_c), take the ratio of the resulting total number of fully (partially) connected networks and all the network realizations for the corresponding pair. We use the same procedure to estimate the average node degree. For clarity purposes the confidence intervals are not depicted in the resulting graphs since they were small—the maximum deviation width was less than 0.6% of the estimated average value for the probability of connectivity or node degree.

Probability of Full Connectivity

Solving (3.17) numerically we calculate the probability of connectivity for different n. Through simulations we obtain the probability of connectivity for the network realizations with respect to the communication radius for different number of nodes deployed and 4 base stations. For comparison purposes, we also simulate the probability of connectivity for a network without sinks—and with four extra nodes to compensate for the lack of base stations.

Fig. 3.6 and 3.7 show the simulation results for small and large n, respectively. Solid lines represent the analytical results of the probabil-

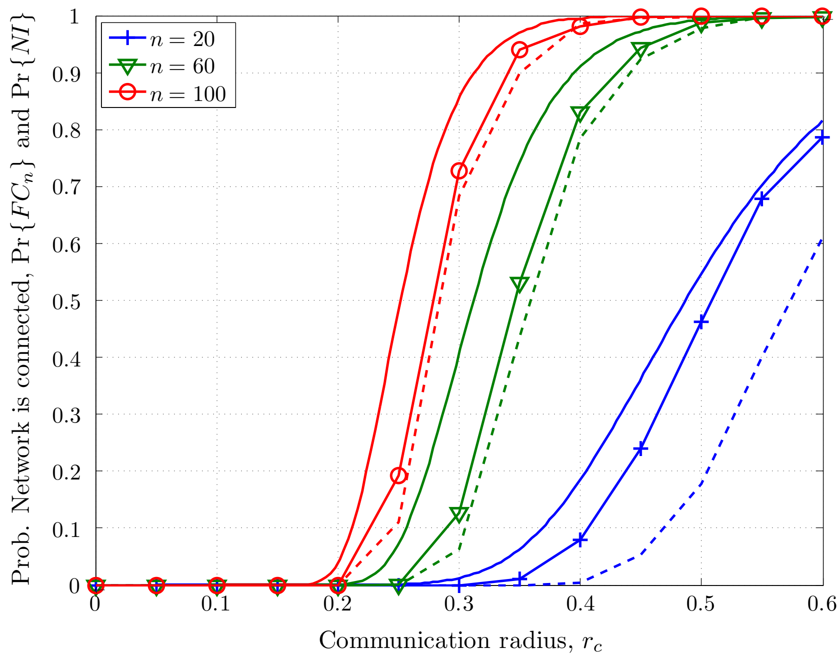

Figure 3.6: Comparison of the probability of having a connected network with and without base stations for small n.

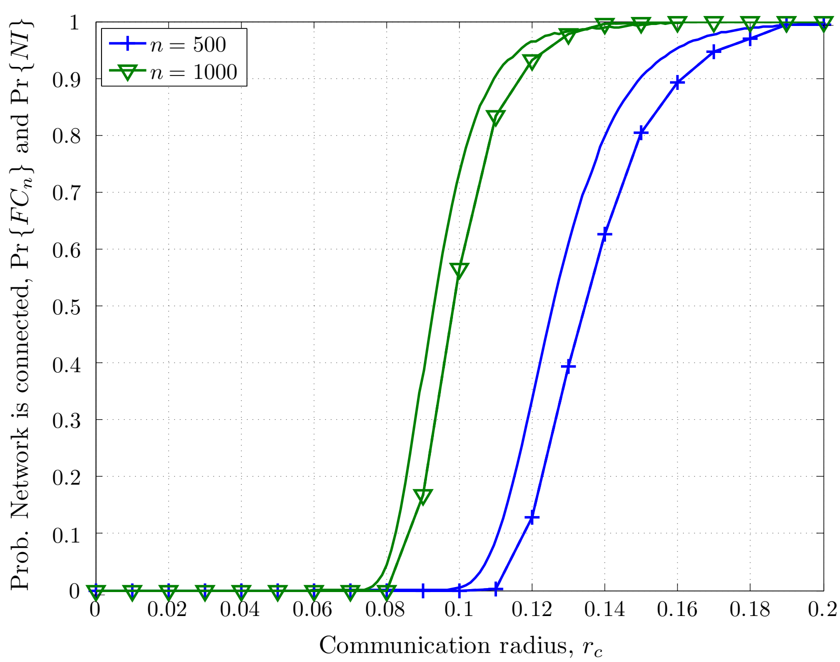

Figure 3.7: Comparison of the probability of having a connected network with and without base stations for large n.

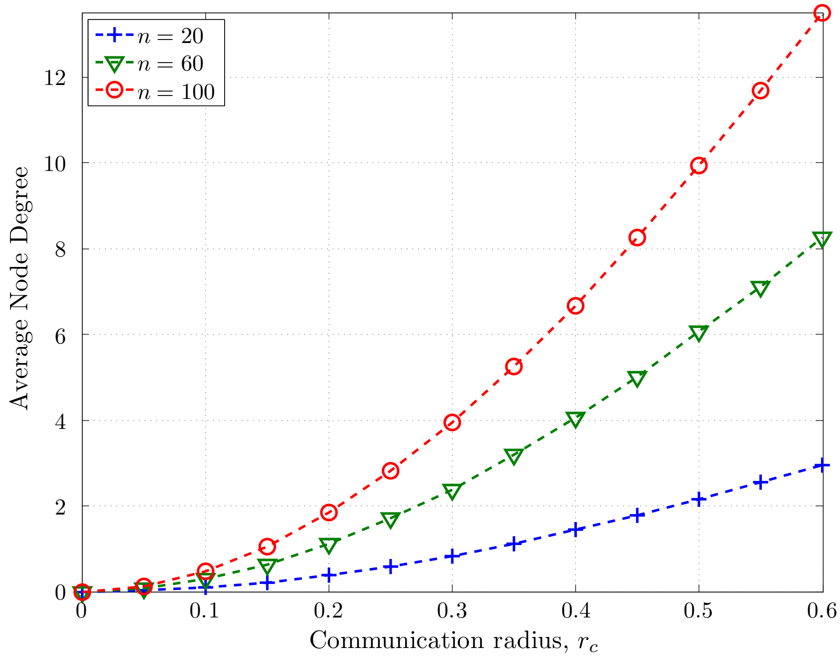

Figure 3.8: Average node degree for small n.

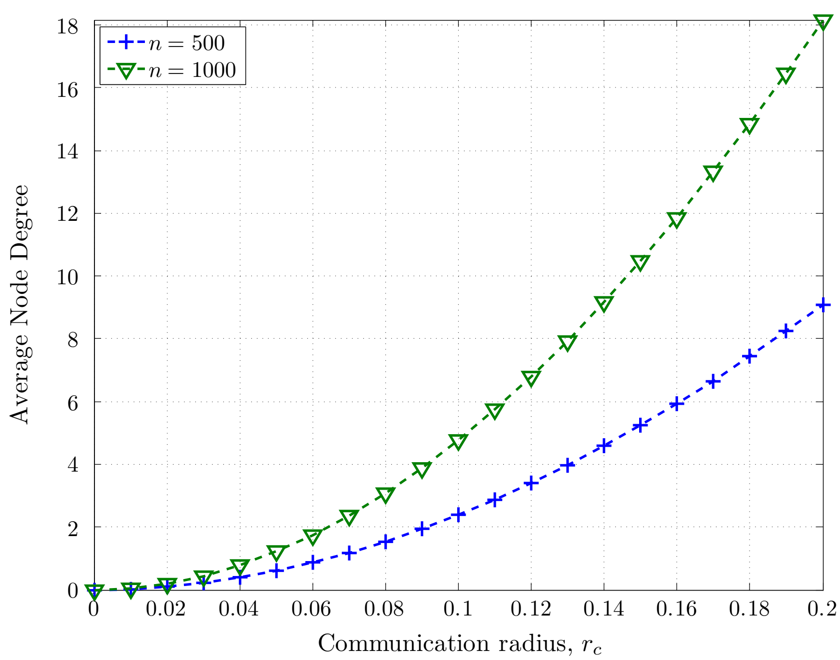

Figure 3.9: Average node degree, large n.

ity of having a connected network. Broken lines indicate the simulation results for networks without sinks. The lines with markers stand for the simulation results of network realizations with sinks. Fig. 3.8 and 3.9 present the results for the average node degrees of the networks.

Probability of Partial Connectivity

Under a setting that allows partial connectivity, Fig. 3.10 and 3.11 show the simulation results for small and large n and 4 sinks. The broken lines are the simulation result of a network that requires full connectivity—the same results as above. The solid lines represent the likelihood of having a partially connected network with respect to the radius of communication, that is the fraction of the nodes that reach a sink. The rest of the nodes belong to isolated components.

Results Discussion

First note that the analytic formula and the simulation results agree in Fig. 3.6 and 3.7. As expected, the probability of connectivity obtained by the analytic results is higher than the actual simulation but the approximation becomes better when n is large or the probability of connectivity is high—since this implies the probability of a node being isolated is small. The discrepancy is larger for small n. Observe how the node degree grows exponentially with respect to the communication radius, hence the potential increase in interference between nodes. When n is large, the number of base stations does not make a difference when con-

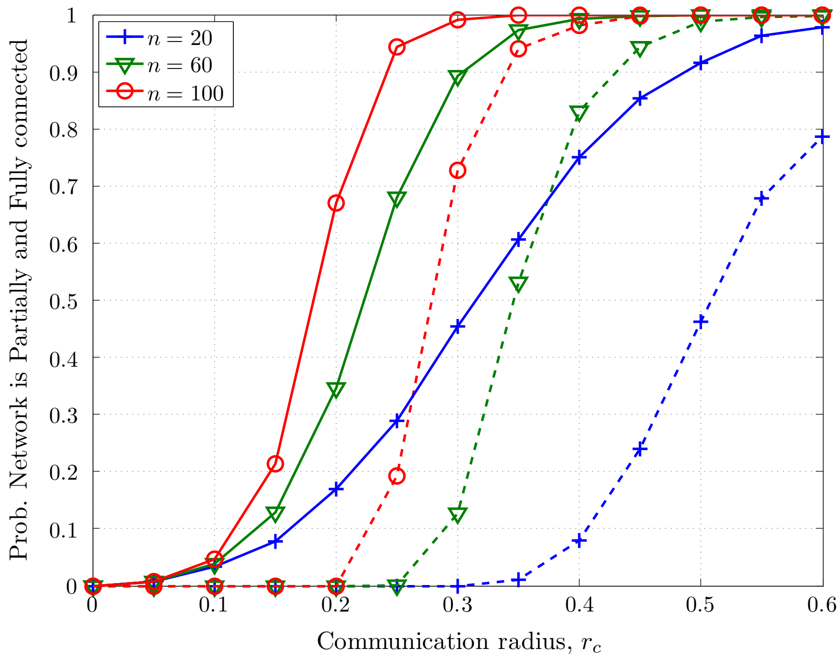

Figure 3.10: Probability of partially and fully connected network, small n.

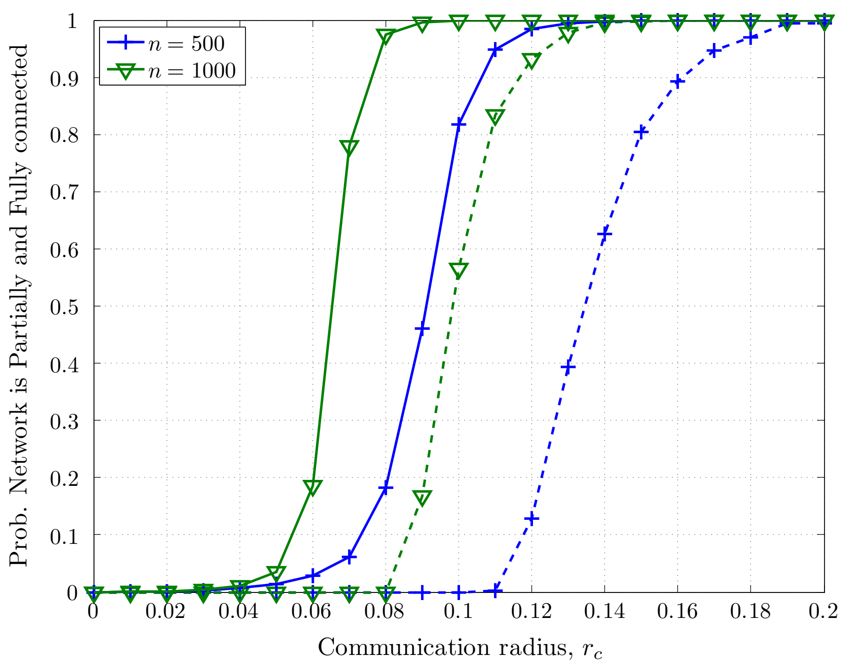

Figure 3.11: Probability of partially and fully connected network, large n.

sidering full connectivity—note how the broken lines match solid ones in Fig. 3.7—but it does for the partial connectivity case.

Now, in Fig. 3.10 and 3.11 note the large increase in the likelihood of connectivity under the partial connectivity scenario—fraction of nodes reaching a sink. This increase comes associated with a lower average node degree, as expected. Having low node degree can improve the network metrics of interference and communication rate. When considering partial connectivity, note that with a relatively small r_c we are able to obtain an almost connected network—one that will have very few isolated nodes, thus saving power in each connected node and extending the network lifetime. As an example, a network with $n = 100$ requires $r_c = 0.30$ to have partial connectivity and $r_c = 0.45$ for full connectivity, thus the average node degree is 4 versus more than 8 in partial and full connectivity, respectively. All these values obtained within the confidence intervals of the stochastic simulations. Summing up, the most isolated node defines the full connectivity of a network [61], and this requires large amounts of resources from the whole network [24].

3.6 Summary

This Chapter presented the connectivity analysis in two-dimensional wireless sensor networks in two settings: full and partial connectivity.

By considering node platforms deployed randomly according to a Poisson process, the connectivity properties of 2-D WSNs were studied. Formulas were obtained for Boolean links as well as deployments under log-normal shadowing communication channels. Stochastic simulations corroborated our analyses and approximations. As expected, the probability of having a connected network is smaller than the probability of not having an isolated node but their difference decreases considerably as the probability threshold increases. Given that in real-world deployments the interest is to have a connected network with very high probability—where the simulations show a very good agreement with the formulae for no isolated nodes—the simple equations (3.7) and (3.10) are relevant to network designers, since they allow them to estimate the minimum node density required when deploying a network randomly.

The second main part of this Chapter considered the full and partial connectivity of 2-D WSNs consisting of a finite number of node platforms, uniformly deployed over a bounded area, along with sinks at fixed locations. An approximate formula was presented which helps to calculate the probability of full connectivity when the number of nodes is large and there are few sinks. The validity of the analysis and its close proximity to network realizations was verified using stochastic simulations. When considering small number of nodes, a slight increase in the probability of having a connected network when using base stations was noted. But more importantly, large improvements of network metrics, like average node degree and power savings by reducing communication radius, were shown when allowing for partial network connectivity.

We have considered the connectivity of WSNs in 1-D and 2-D. In the following Chapter we present an application of such results. This helps to illustrate their use and relevance when designing WSNs.

CHAPTER 4

TOPOLOGY CONTROL OF CONNECTED WIRELESS SENSOR NETWORKS

4.1 Introduction

Wireless sensor networks (WSNs) are composed of nodes having processing, sensing, and radio capabilities along with a constrained power supply. Due to the energy scarcity of the nodes, using specialized mechanisms to extend the lifetime of a WSN is an important topic and a key design challenge [46]. A possible method to achieve this lifetime extension is by controlling the network topology through a sleeping scheme.

In the previous Chapters we considered connectivity in 1-D and 2-D WSNs. This Chapter addresses the problem of improving lifetime of a WSN by controlling its topology through the use of information about data correlation—data obtained by sensing the physical environment. With this information the nodes are able to create a two-tier network and thus extend the functional lifetime of the WSN. The objective of this scheme is to increase the lifetime of a WSN while keeping its connectivity.

An example of an application where it is crucial to use the available energy in the most efficient way is the use of WSN for monitoring physical structures. Practical implementation of this monitoring systems include SCADA systems, water supply pipes, and Structural Health Monitoring (SHM) in general [53]. Note that, among other type of measurements, WSNs are used for modal vibration in SHM, a type of monitoring where a strong spatial correlation in the sensed data exists [53]. Relative to other general applications—like military operation or commercial sensing—monitoring physical structures requires wireless networks with extremely long lifetime. In addition, since node platforms are generally embedded into the structures it may be impractical or impossible to install more platforms or to recharge their power supplies.

This Chapter presents a topology control mechanism for highly redundant WSNs as a way to improve the network functional lifetime. In addition, we analyze the relation between the increase in lifetime and the connectivity properties of such topology control mechanism. It is possible to extend the functional lifetime of a network by controlling its topology through a scheme that sets nodes to sleep, though at the expense of performance loss in other metrics. But even when controlling its topology, for a WSN to carry out its functions it should remain connected at all times.

In particular, we consider a two-dimensional (2-D) network where nodes are deployed randomly following a Poisson point process. The method to address the lifetime extension problem is by using the theory of point processes along with results from random geometric graphs.

When designing a random WSN for a given connectivity level, there is a trade-off between network reliability and energy efficiency [88]. Given the base network lifetime and its desired extension factor, this Chapter provides formulas for the required density of nodes to be deployed in order to keep the network connected with a given probability. Using the proposed topology control mechanism, and by adjusting the density of the node deployment, the lifetime of a network can be extended in a simple, autonomous, and scalable way. The main idea behind our topology control scheme is to utilize the information obtained by the nodes from the environment where they are deployed. That is, the strategy is to take advantage of data correlation to control the topology of a WSN.

This Chapter is organized as follows. Next Section 4.2 presents the related work in the literature of the topology control and network lifetime extension. Then Section 4.3 states the general problem to be addressed along its assumptions. Section 4.4 describes the topology control mechanism, including its advantages and shortcomings compared to other similar schemes. Finally, Section 4.5 contains the mathematical model and analysis for the lifetime extension, including the simulation results for network deployments under models assuming Boolean communication links and log-normal shadowing in the communication channel.

4.2 Related Work

While there are several schemes in the literature that address sleep based topology control in WSNs, there is no mechanism that exploits the information provided by the correlation among sensed data [78].

The following research is related to this Chapter. With respect to the analysis of network lifetime extension through clustering, the authors in [41] propose a protocol for WSNs called LEACH that uses clusters to form a two-tier network to reduce energy consumption. The clusters are formed in a random way, that is without using information from the physical phenomena under sensing. Instead the purpose of the protocol is to have certain average quantities of cluster heads at any given time while using efficiently energy and bandwidth in wireless communications. The impact of spatial correlation on routing for structured network deployments is analyzed in [60]. There, the authors find the optimal size of a cluster using the information of joint entropy for a set of sources. They consider the entropy as a function of the distance between nodes and the research shows that there is a near-optimal cluster size, in terms of energy efficiency, that performs well over a wide range of spatial correlations.

Data-aggregation algorithms are also useful to extend the functional lifetime of a WSN. By performing in-network processing through data aggregation, node platforms are able to reduce the amount of information transmitted, thus saving energy. These algorithms can operate as a complement to sleep-based topology control techniques in order to increase network lifetime. In essence, data aggregation is a routing scheme with compression of correlated data. The authors in [69] show the strong relation between routing and source coding.

The current approaches of sleep-based topology control for WSNs depend on design assumptions and goals. A way to maximize the functional lifetime of sensor networks is presented in [34]. The authors focus on maximizing, through a communications scheme, the number of data that is transmitted from each node to one sink.

The work in [83] presents the algorithm GAF whose purpose is to reduce energy consumption by setting redundant nodes, from a routing perspective, to sleep. GAF requires the nodes to generate a grid, thus nodes require location information.

To extend the network lifetime, techniques like the asynchronous, energy efficient scheduling are relevant. In the randomized independent sleeping RIS scheme node platforms are either sleeping or active at every time slot with certain probability [47]. Thus the probability of being active is a parameter that controls the factor of increase of network lifetime. The (RIS) scheme calculates the number of sensors required to have a k-coverage network on a unit square area. The work in [87] presents a scheme using combinatorial designs. This technique assures conditions where a particular structure exists in the active nodes, allowing each node to keep independent sleeping cycles.

4.3 Model and Problem Formulation

Our goal is to extend the functional lifetime of a WSN.

Definition 4.1. *Base Network Lifetime: Consider as the* base *network lifetime the maximum life period of a continuously active node platform.*

Then given this base lifetime and its desired extension factor, we will provide formulas for the required density of nodes to keep the network connected with a given probability under the topology control scheme. The concept of lifetime is in the sense of functional lifetime for WSNs [34]:

Definition 4.2. *Network Functional Lifetime: The maximum time a certain network metric can be kept as designed.*

In particular, we consider the metric of network connectivity. Then network functional lifetime will be the time when the network becomes disconnected.

Considering the connectivity and topology control of a WSN, the relevant design questions are:

1. how to form clusters of nodes in an autonomous way using only local information while keeping network connectivity?
2. how to provide an energy-efficient scheduling mechanisms in order to increase network lifetime?

The heuristics behind a possible scheme is to allow nodes to evaluate data correlation within their neighborhood and then form clusters. After that, some nodes inside the cluster decide to go to backup mode (or sleep), while others remain in active mode.

To investigate our scheme we make the following assumptions. Consider a random deployment of sensor platforms in a unit area square according to a homogeneous spatial Poisson process Π with given intensity. The solution approach would be the same for a square with different area, just with proper scaling factors. Let the node platforms be static, each with fix transmission power. Suppose that the communication links model is Boolean such that each node has a fixed communication radius r_c. For an illustration see Fig. 4.1.

Assume the network has a high density of nodes such that, at deployment time, it is connected with high probability—more details about

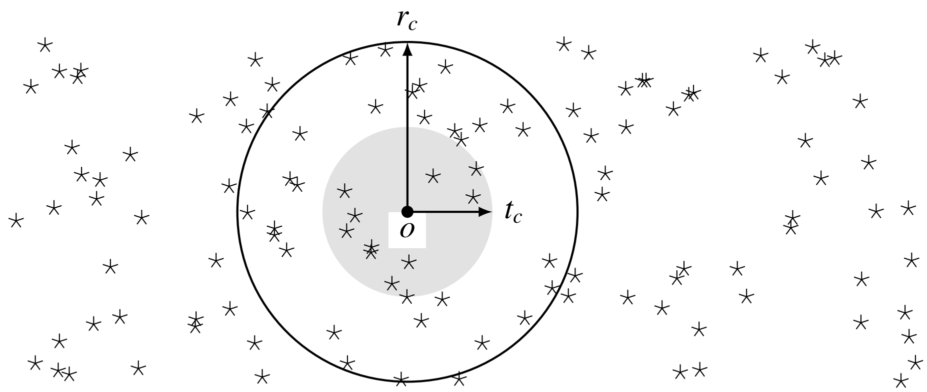

Figure 4.1: Example of a realization of a network. Node o has communication radius r_c, when $\sigma = 0$. The radius where there is perfect correlation is t_c.

this in Section 4.5. Further, assume that nodes are able to measure the distance between them, for example using received signal strength indicator (RSSI) [46], a widely available capability in off-the-shelf node platforms. Finally assume that nodes are able to determine data correlation information—this is a standard assumption in research literature although it is still a challenging problem in practical setups.

Justifications of our assumptions are in order. To ponder a more tangible setting consider a WSN that will be used for monitoring physical structures. In this situation, the node platforms will be attached to a structure, hence the resulting WSN will be static. The Boolean communication links model and the fixed communication radius allows us to isolate the parameter of interest, that is network functional lifetime. Moreover there are ways to account for the irregularity in the quality of wireless communication links, for example by estimating their quality and compensating for routing decisions [84]. The high node platform density allows for node redundancy, something specially desired in WSNs since extra nodes boost network lifetime.

The rationale behind the distribution of nodes as a spatial Poisson process is that this point process is the most random way to describe a deployment. That is, if we consider disjoint spatial areas containing nodes, then those areas are statistically independent and the Poisson model follows. Regarding the model for spatially varying phenomena that the WSN is monitoring, we assume a Boolean one. Thus if two nodes are within a distance t_c, then they are perfectly correlated, see Fig. 4.1. The value of the parameter t_c depends on the particular application of the WSN. Another possible model for the spatially varying phenomena could be a Gaussian random fields, but since our scheme requires either perfect correlation—above some threshold—or not, then the Boolean model is more appropriate.

4.4 Topology Control Mechanism

As we mentioned, a way to accomplish lifetime extension in a WSN is to have redundant deployment of node platforms along with a sleeping scheme. Other methods comprise energy efficient algorithms, including, for example, medium access control, low power listening, data aggregation or in-network processing, and energy efficient scheduling.

Due to their relative simplicity, sleep-based topology control is a popular and important technique used in WSNs to reduce the energy consumption of the individual node platforms and, as a consequence, sleep-based topology control is useful to increase the functional lifetime of the networks. In general, topology control mechanisms are distributed and decentralized, thus scalable and they can either (1) adjust the transmission power of each node; or (2) use sleep cycles of the nodes. A comprehensive survey for the former method is provided in [68].

The topology control scheme that we present is useful for applications where precise positioning of a node platform is not essential and nodes are deployed with high redundancy. Platform redundancy may occur due to unfeasibility to deploy more nodes—as a result of remoteness of location, extreme weather or simply because nodes are embedded into infrastructure.

While in the literature there are several schemes that address sleep-based topology control in WSNs, there is no mechanism that operates exploiting the information provided by the correlation among sensed data. This Section describes a method to control the topology of a WSN using information about data correlation. With the correlation information, the node platforms are able to create a two-tier network—one tier of active while the other of backup nodes—and thus extend the functional lifetime of the WSN.

Basically, as described below, the mechanism consists of two steps:

1. autonomous creation of clusters (§4.4.1)
2. use of a scheduling algorithm within these clusters (§4.4.2)

This idea is illustrated in Fig. 4.2. Our scheme has practical relevance since it is simple, localized, decentralized, and scalable. Moreover, using

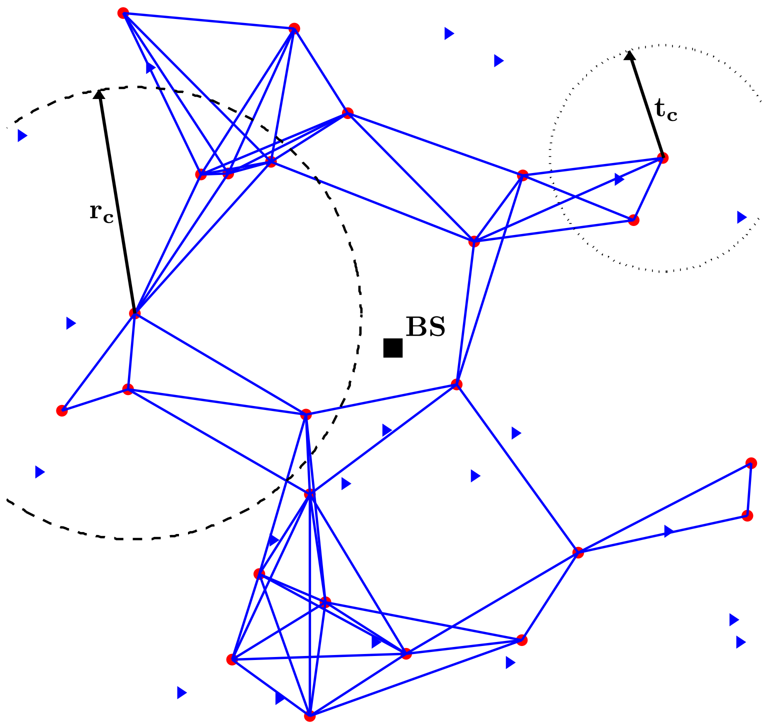

Figure 4.2: Basic idea of the scheme where nodes are randomly deployed. Solid circles represent active nodes and triangles represent backup nodes.

correlation information allows relaxing common assumptions imposed in other topology control schemes.

4.4.1 Cluster Formation

A brief explanation. In this step there is a combination of the information obtained by the correlation of the sensed data that helps creating clusters of node platforms.

The topology control scheme forms clusters by using the degree of correlation of the sensed data. This can be used in applications where the measurement data of physically close node platforms exhibit high degree of correlation. For example, the spatial correlation concept is also used implicitly in data fusion algorithms, where it allows them to eliminate redundancy of the information being relayed [46]. We can consider these systems to be physically-integrated in the sense that the formation of clusters depends on the underlying physical phenomena. Below we will assume a model of a physical environment with properties of well-defined zones of spatial correlation.

Next, we provide more details on the clustering step. As mentioned above, we use the principle of spatial correlation—similar to [60] where the entropy-based measurement is a function of distance. Per our assumption of a Boolean model for correlated data, the platforms that are close to each other will have their readings correlated and each node will possess the correlation measurement of its immediate neighbor.

When the network is deployed all the nodes are active and their initial task is to discover their immediate neighbors. To have a localized and autonomous scheme, at time of deployment each node becomes either an *active* node or a *backup* one according to a tuning parameter, based on the desired network structure. The active nodes act as cluster heads, meaning that they form virtual clusters using the backup nodes with whom they have the highest data correlation that is above certain threshold—spatially located within distance t_c.

Thus clusters are created autonomously and at random positions, depending on the location of the active nodes. Note that this method allows for backup nodes to belong to multiple clusters at the same time. Although it would be possible to force the backup nodes to belong to only one cluster—for example using a conflict resolution mechanism.

Once clusters are formed the challenging question is how to guarantee a connected network. As discussed in the previous Chapters, a network is connected with high probability if for a given area the node platforms density is greater than a critical density. Then, with the theory developed in the previous Chapters, we can estimate the node density required to achieve an almost sure connectivity using probabilistic values—for example having a connected network with 99% of chance.

4.4.2 Sleeping Scheduling

A brief explanation. Once clusters are formed, a scheduling step is implemented to save energy of the node platforms. That is, within the clusters of nodes an energy efficient random scheduling mechanism controls the underlying topology of the WSN.

Next we provide more details on the scheduling step. After the clusters are formed some nodes within them can go to sleep, thus increasing the network functional lifetime. The nodes under the sleeping state wake-up with a very low duty-cycle and verify that there is still an active node within its cluster. An asynchronous wake-up mechanisms, for example, can be used to implement this scheduling.

The way in which active and backup nodes are selected indirectly determines, in an automatic way, which nodes will be sleeping. Reflecting on the principles used in the RIS scheme [47], backup nodes belonging to a cluster go to sleep randomly and independently from each other. Therefore, a random energy efficient scheduling mechanism [78] will help on the energy savings of the node platforms.

Given that the communication radius of the node platforms is much larger than the distance where two nodes are strongly correlated, and that the network is connected, it is possible to select a representative node from a cluster of nodes. In this case we propose just to select the active node, but it is possible to implement an algorithm that takes other factors into consideration, like node degree. To provide for network robustness, the backup nodes wake up periodically, although with extremely low-duty cycle. Let us call this brief activity period *probing*.

This energy saving mechanism is not optimal, but since it is random it does not require complete information about the state of the network. The backup node duty cycle is dictated by the underlying application of the WSN and its objective is to verify if an active node malfunctions. In this manner, using a random timer, a backup node wakes up and checks for cluster head activity—for example through a HELLO message. If the cluster head is still alive, the backup node goes back to sleep, otherwise it will stay in active mode.

4.4.3 Advantages and Drawbacks

The main advantage of the topology control mechanism proposed in this Section is the extension of WSNs lifetime. Additionally, the scheme is localized and information about node platform location is not required.

Moreover the creation of clusters requires information only of the immediate neighbors and, as a result, is done in a distributed manner, hence the scheme is scalable.

The principal shortcoming of the mechanism is that it is applicable under environments where the underlying physical phenomenon has high spatial correlation. Also, note that each node has to obtain the level of data correlation with its neighbors—though this is a common assumption in the literature [60].

Something that can be considered as a drawback of our scheme is that it does not provide for a uniform depletion of energy. Although depending on the application, like in Structural Health Monitoring (SHM), the uniform energy depletion is a secondary concern, while the primary matter is extending the lifetime of the network and having nodes continuously sensing for unpredictable events.

4.4.4 Specialization

In order to make the analysis in the following Section concrete let us define specific characteristics of an environment under sensing and the corresponding topology control scheme implementation.

We assume that the physical phenomenon under monitoring has a high degree spatial correlation, such that a perfect correlation exists between two nodes if they are close enough, similar as the model used in [60]. Also, we consider a homogeneous energy consumption rate per active node, such that it is negligible when a node is sleeping.

To form clusters a random scheme is used, where, at time of deployment, nodes become either active or backup with probability p_a and p_b respectively, $p_a + p_b = 1$. This random assignment of backup (sleeping) nodes is the basis of the RIS mechanism [47]. The value of p_a represents a tuning parameter that controls the factor of the lifetime extension.

By using this simple scheme specialization we are able to provide explicit formulas in the next Section. The network topology is controlled by the backup nodes that go into sleeping state and wake-up with a low duty-cycle, becoming active nodes if needed.

4.5 Network Lifetime Extension Analysis

In this Section we analyze the network functional lifetime extension. We use the topology control scheme developed in the previous Section, assuming the specialization described in §4.4.4.

According to the particulars of the model assumptions, there are several definitions of lifetime of a WSN in the literature. It can indicate the time when either:

- the first node depletes its energy, or
- there is a certain percentage of nodes alive, or
- the network can still carry out useful operations.

Given that the last definition encompasses a state where a WSN can still function—although working at a lower quality of service, the third definition seems the more adequate for most of the practical deployments. Then in this work we consider the functional lifetime definition with the level of network connectivity as a performance metric.

Let the desired network lifetime extension factor (EF) be $EF = k$, $k \in \mathbb{R}^+$. Denote by S the field of interest and, without loss of generality, let $S = (0,1) \times (0,1)$, a unit area square. Consider N to be a random variable representing the number of node platforms deployed in S. Assume N is distributed according to a homogeneous Poisson process with intensity ϱ, denote it by Π_ϱ. Then $N(A)$, for $A \subseteq S$, has the Poisson distribution with parameter ϱ, designated by $\mathscr{P}(\varrho)$, and ϱ represents the node density of the network deployment. Moreover, assume that ϱ is at least as large as the density required for having the probability of connectivity to be $\Pr\{C\} = p$ for a network in S.

Let p_a and $p_b = 1 - p_a$ be the probabilities that a node becomes active or backup, respectively, at deployment time. Hence, we have generated a marked Poisson process [45, Chapter 5], where there are two independent random variables with distribution $N_a \sim \Pi(\varrho p_a)$, $N_b \sim \Pi(\varrho p_b)$, representing the number of active and backup nodes respectively. Since $N(A)$ is the total number of nodes in set $A \subseteq S$, let $N_a(A)$ and $N_b(A)$ be the number of active and backup nodes respectively in A. By properties of homogeneous marked Poisson spatial processes, we have $N(A) = N_a(A) + N_b(A)$.

Denote by $\alpha = p_a\varrho$ and $\beta = p_b\varrho$. Since $p_a + p_b = 1$, then $\alpha + \beta = \varrho$. Given that α and β are assigned randomly, Π_ϱ represents a marked Poisson process, composed of two independent Poisson processes, denote them by Π_α and Π_β. Let us analyze the properties of a point at o. Recall that by the Slivnyak's theorem we have Π_β has the same distribution as $\Pi_\beta \cup \{o\}$. Then it suffices to analyze the point at o to obtain formulas for the rest of the points in the network realization.

By using the topology control described in §4.4.4, the desired lifetime extension factor EF represents the average number of neighbors of o within $\mathrm{b}(\boldsymbol{x}_o, t_c)$. Where $\mathrm{b}(\boldsymbol{x}, r)$ is a disc with center $\boldsymbol{x} \in S$ and radius r. Considering the Poisson process Π_β and the perfect spatial correlation radius t_c, we have

$$\mathbb{E}\{\, N(\mathrm{b}(\boldsymbol{x}_o, t_c)) \,\} = \beta |\mathrm{b}(\boldsymbol{x}_o, t_c)| \tag{4.1}$$

where $|B|$ represents the area of B.

Then for $EF = k$ we have

$$\beta = \frac{k}{|\mathrm{b}(\boldsymbol{x}_o, t_c)|} \tag{4.2}$$

To provide an insight about the behavior of the number of neighbors of o as a function of t_c, consider the distribution of its nearest neighbors. Take the network realization Π_β, label the j^{th} nearest neighbor of o by j and let $D_j(r)$ be the distribution function of the j^{th} nearest neighbor, that is $\Pr\{\, N(\mathrm{b}(\boldsymbol{x}_o, r)) \leq j \,\}$. Denote by l_j the distance from o to j. Note that $l_1 > r$ if and only if $\mathrm{b}(\boldsymbol{x}_o, r)$ has no node, then

$$\Pr\{\, l_1 > r \,\} = \exp(-\beta \pi r^2)$$

So it follows that

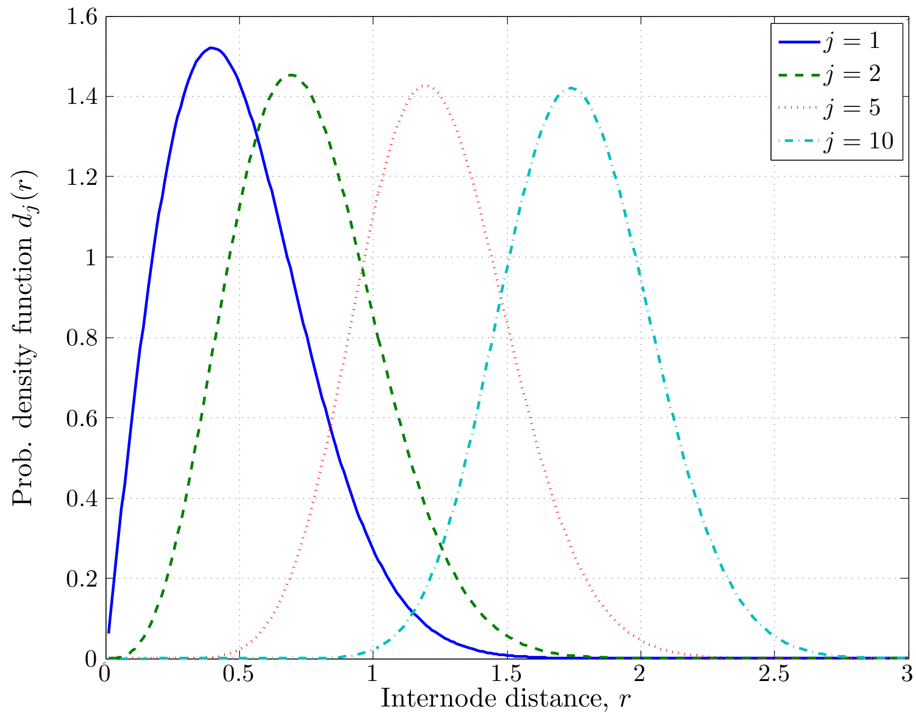

Figure 4.3: Probability density function of $d_j(r)$ for different j.

$$D_1(r) = 1 - \mathrm{e}^{-\beta\pi r^2} \tag{4.3}$$

From it, the corresponding density function is

$$d_1(r) = 2\beta\pi r \mathrm{e}^{-\beta\pi r^2} \tag{4.4}$$

Note that $D_1(r)$ implies that the area $A_1 = \mathrm{b}(\boldsymbol{x}_o, l_1)$ is exponentially distributed with parameter β. Let $A_j = \mathrm{b}(\boldsymbol{x}_o, l_j)$, it follows that $A_1, A_2 - A_1, \ldots$ are independent and exponentially distributed random variables with parameter β. Hence $D_j(r)$ has a Gamma distribution with density

$$d_j(r) = \frac{2(\beta\pi r^2)^j}{r(j-1)!} e^{-\beta\pi r^2} \qquad (4.5)$$

Fig. 4.3 shows $d_j(r)$ for different j. It can be shown that the maximum of (4.5) occurs when

$$r = \sqrt{\frac{j-1/2}{\pi\lambda}} \quad , \qquad \forall j$$

This implies that when λ increases, l_j decreases, as expected.

Now, given β from (4.2), we need to obtain the rest of the parameters for the topology control scheme. With S, r_c, and the desired p for $\Pr\{C\}$, to obtain α we just need to use the equations from the previous Chapter—either (3.7) or (3.10) according to the channel model under analysis—in order to obtain the required node density for the given connectivity probability. Finally, since $\varrho = \alpha + \beta$, we get $p_a = \alpha/\varrho$ and $p_b = \beta/\varrho$. Hence we get all the parameters for the topology control scheme design.

4.6 Simulation Results

This Section presents simulation results. Take S as a unit area square where nodes are randomly deployed according to $\Pi(\varrho)$, with ϱ varying from 20 to 60. Let the communication radius be $r_c = 0.45$ and the correlation radius be $t_c = 0.15$. The procedure for the simulations is by using the Monte Carlo method with 10^4 random replications for each set of parameters. Table 4.1 shows the simulation parameters. For clarity pur-

Table 4.1: Parameters used in the simulation setup.

Parameter	**Value**
reps	10^4
S	$(0,1) \times (0,1)$
ϱ	{20, 30, 60}
r_c	0.45
t_c	0.15

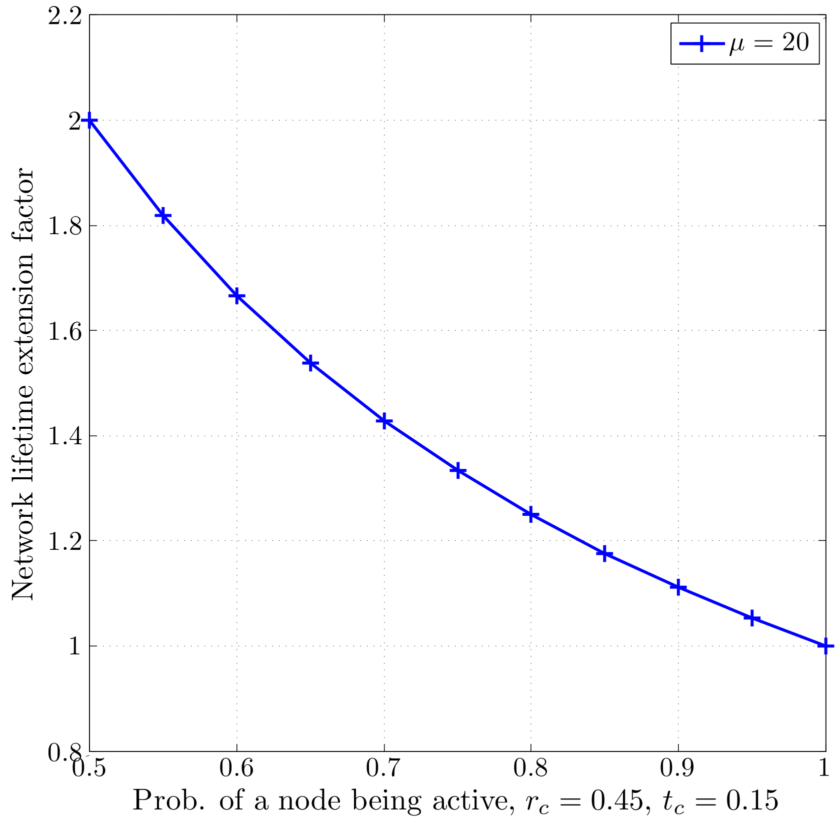

Figure 4.4: Network Lifetime Extension Factor for a given node density and network parameters.

poses the confidence intervals are not depicted in the resulting graphs, though we remark that they were small—the maximum deviation width was less than 0.4% of the estimated average value.

Fig. 4.4 shows the network lifetime extension factor with respect to the probability of a node becoming active at deployment time for a sample deployment density. As expected the network lifetime doubles when the probability of a node being active at deployment time is one half.

4.7 Summary

This Chapter presented a topology control scheme for WSNs with the aim of increasing the network functional lifetime. In addition, we analyzed the lifetime extension and connectivity properties of such mechanism for two-dimensional WSNs.

We considered node platforms deployed randomly and obtained formulas of the network lifetime extension for nodes having Boolean communication links. It was shown that, when choosing the right density of node platforms to be deployed, the topology control mechanism will be able to extend the network lifetime for the required factor while keeping the WSN connected.

The fundamental idea of the topology control scheme is to use the information about the correlation of the data obtained by sensing a physical phenomenon to control the individual node platforms composing the WSN. This scheme was shown to increase the functional lifetime of the WSN; task achieved by reducing the energy consumption of the node platforms through an energy-efficient scheduling mechanism. The scheme offers a practical way to define clusters of nodes while keeping the network connected. In particular, the topology control mechanism consists of two steps: (1) autonomous creation of clusters and (2) use of a scheduling algorithm within these clusters. We presented the analysis of a tractable model of this scheme, showing its performance along with simulations that support the analytical results. The scheme for lifetime extension has practical relevance since it is simple, localized, decentralized, and scalable.

In this dissertation we covered the connectivity of WSNs in 1-D and 2-D. In addition and through an application, in this Chapter we illustrated how to use our results and their relevance when designing WSNs. The next Chapter presents the conclusions of our work.

CHAPTER 5

CONCLUSIONS

This dissertation presented an analysis of a topological characteristic of finite random wireless sensor networks (WSNs). In particular, it focused on the connectivity of one dimensional (1-D) and two dimensional (2-D) random network deployments. In addition, a simple mechanism was developed for controlling the topology of a 2-D WSN. This work provided formulas for the probability of having a connected wireless sensor network under different settings, including variations in the communication links, channel randomness, and network infrastructure. We also presented formulas that estimate the increase in network functional lifetime when using a mechanism for autonomously controlling the topology of a random WSN.

A brief introduction of WSNs was presented in Chapter 1, including their characteristics and design challenges. That Chapter outlines the problems that were analyzed in this dissertation along with the techniques used to address those problems. In addition, Chapter 1 lists our research that is related to this dissertation.

The problem of network connectivity in 1-D random deployments with finite number of node platforms was considered in Chapter 2. To mathematically analyze the networks, we abstracted the WSNs as random graphs and proposed a communication model for deterministic and random channels. By using combinatorics and stochastic geometry theories, we derived equations on the probability of connectivity of such random WSNs, while assuming networks with and without infrastructure—in the form of base stations. We noted how the requirements for having a connected WSN may be relaxed since there can be more than one connected component of a network. In addition, we also provided a general formula for the probability of connectivity when the node platforms communication radii are random—including a closed form formula for the particular case when the communication radii follow a uniform distribution. In this setting, we noted a slight increase in the probability of network connectivity for given parameters, that is the requirements on the number of nodes to have a connected WSN may be relaxed when such nodes have random communication radii. Furthermore, the random communication radii model can represent certain randomness in the communication channel.

The network connectivity in 2-D finite random deployments was analyzed in Chapter 3. We considered the problem of connectivity in random deployments of nodes with and without infrastructure. Again, we abstracted the WSNs as random graphs with a particular model for the communication links. We analyzed the situations for deterministic and random communication channels and, by using random geometric graphs and stochastic geometry theories, derived equations on the probability of connectivity. After that, we extended the analysis for the case when there are multiple base stations in the random deployment. Additionally, we considered the effects on WSNs metrics by allowing for partial network connectivity. We showed the benefits of using more than one sink as well as how partial connectivity with sinks improves network lifetime, average node degree, and communication througthput.

A simple topology control scheme that extends the network functional lifetime of a 2-D random WSN was presented in Chapter 4. That Chapter considered the problem of network lifetime extension in 2-D network deployments after the network has been setup. We abstracted the WSN as a random graph with a particular model for the communication channel—assuming deterministic communication links or Boolean, as well as for the field of interest underlying physical phenomena. By using stochastic geometry theory, we derived the formulas on the functional lifetime extension gained by using the topology control scheme. This mechanism operates by exploiting the information provided by the correlation among sensed data, thus allowing the autonomous creation of clusters in a simple, localized, decentralized, and scalable way.

Comparing with asymptotic approximations, the main advantage of using the exact connectivity formulas presented in this work for the analysis and design of random wireless sensor networks with finite number of node platforms is the accuracy enhancement of the probability of connectivity estimate. In addition, the exact formulas present a reasonable performance improvement without excessive overhead in the calculations of the probability of connectivity, when compared to the asymptotic approximations. The principal reason to obtain exact formulas instead of using only stochastic simulations for the design of a random WSNs is the computational scalability issue: stochastic simulations of such networks do not scale linearly with the number of node platforms being simulated [10, 81].

A practical application of this theoretical work is to monitor physical structures. These structures are pervasive—from museums, data centers, water supply systems, and bridges to farms, mines, and islands. When designing a random WSN for such settings it is critical to have the network connected in order to be able to extract the data. Our work provides a theoretical framework for some of those practical deployments.

5.1 Possible Extensions of This Work

In this work we showed exact analyses assuming well-defined probability distributions for the deployment of the node platforms. Experimental data would be useful to obtain more accurate probability distribution of the platforms on real world deployments. Moreover, all the connectivity analyses presented in this dissertation have been done in a theoretical setting. It is of interest to obtain practical results. The same holds true for an implementation of the topology control scheme in large-scale, high-density, WSNs deployments.

Additional future extensions to this dissertation include the following. Addressing connectivity in 2-D networks with finite number of node platforms and with communication interference in addition to log-normal shadowing, while obtaining mathematically rigorous results. Also, it may be fruitful to model the interdependence of the log-normal shadowing parameters—path-loss and variance of the underlying normal random variable—and possibly include their temporal variation. What is more, it would be interesting to obtain formulas for the optimal location of the base stations for a given network infrastructure setup for both 1-D and 2-D deployments. In addition, it would be of practical relevance to address the trade-off costs between having a WSN with and without infrastructure. Further research is necessary to find an optimum balance between cost and performance of a wireless network deployment.

Finally, notice that we did not address any temporal analysis, so dealing with connectivity of WSNs while considering its behavior as a function of time, instantaneously, may appear that the connectivity increases under high variance in the log-normal shadowing model. But, on average, the probability of network connectivity might be the same as without considering the log-normal shadowing effect. It may be worthwhile to carry out such analysis, since this would improve the understanding, modeling, and design of future wireless sensor networks.

REFERENCES

[1] N. Ahmed *et al.*, “The holes problem in wireless sensor networks: A survey,” *ACM Mobile Comput. and Commun. Review*, vol. 9, no. 2, pp. 4–18, Apr. 2005.

[2] S. Asmussen and P. W. Glynn, *Stochastic Simulation: Algorithms and Analysis*, 1st ed., B. Rozovskii and G. G., Eds. New York, NY, USA: Springer, 2007.

[3] X. Bai *et al.*, “Deploying wireless sensors to achieve both coverage and connectivity,” in *Proc. of the 7th ACM international symposium on Mobile ad hoc networking and computing (MobiHoc’06)*, Florence, Italy, May 22–25, 2006, pp. 131–142.

[4] R. Béjar *et al.*, “Sensor networks and distributed csp: communication, computation and complexity,” *Artificial Intelligence*, vol. 161, no. 1–2, pp. 117–147, Jan. 2005.

[5] S. A. Bermúdez and S. B. Wicker, “On the connectivity of finite wireless networks with multiple base stations,” in *Proc. 17th International Conference on Computer Communications and Networks (ICCCN’08)*, Saint Thomas, U.S. Virgin Islands, Aug. 3–7, 2008, pp. 1–6.

[6] ——, “Taking advantage of data correlation to control the topology of wireless sensor networks,” in *Proc. 15th International Conference on Telecommunications (ICT’08)*, Saint Petersburg, Russia, Jun. 16–19, 2008, pp. 1–6.

[7] ——, “Connectivity analysis under shadowing in wireless networks with data correlation topology control,” in *Proc. IEEE International Symposium on Wireless Pervasive Computing (ISWPC’09)*, Melbourne, Australia, Feb. 11–13, 2009, pp. 1–5.

[8] ——, “Connectivity of finite wireless networks with random communication range nodes,” in *Proc. IEEE International Conference on Communications (ICC’09)*, Dresden, Germany, Jun. 14–18, 2009, pp. 1–5.

[9] ——, “Partial connectivity of multi-hop two-dimensional finite hybrid wireless networks,” in *Proc. IEEE International Conference on*

Communications (ICC'10), Cape Town, South Africa, May 23–27, 2010, pp. 1–5.

[10] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. of the 3rd ACM international symposium on Mobile ad hoc networking and computing (MobiHoc'02)*, Lausanne, Switzerland, Jun. 9–11, 2002, pp. 80–91.

[11] ——, "On the connectivity of *Ad Hoc* networks," *The Computer Journal*, vol. 47, no. 4, pp. 432–447, 2004.

[12] C. Bettstetter and C. Hartmann, "Connectivity of wireless multihop networks in a shadow fading environment," *Wireless Networks*, vol. 11, no. 5, pp. 571–579, Sep. 2005.

[13] C. Bettstetter *et al.*, "How does randomized beamforming improve the connectivity of ad hoc networks?" in *Proc. IEEE International Conference on Communications (ICC'05)*, vol. 5, Seoul, Korea, May 16–20, 2005, pp. 3380–3385.

[14] P. Billingsley, *Probability and Measure*, 3rd ed. New York, NY, USA: John Wiley & Sons, 1995.

[15] B. Bollobás, *Random graphs*, 2nd ed. New York, NY, USA: Cambridge University Press, 2001.

[16] L. Booth *et al.*, "Covering algorithms, continuum percolation and the geometry of wireless networks," *Ann. Appl. Probab.*, vol. 13, no. 2, pp. 722–741, 2003.

[17] Y.-C. Cheng and T. G. Robertazzi, "Critical connectivity phenomena in multihop radio models," *IEEE Transactions on Communications*, vol. 37, no. 7, pp. 770–777, Jul. 1989.

[18] K. Chintalapudi *et al.*, "Structural damage detection and localization using NETSHM," in *Proc. of the 5th international symposium on Information processing in sensor networks (IPSN'06)*, Nashville, TN, USA, Apr. 19–21, 2006, pp. 475–482.

[19] D. R. Cox and V. Isham, *Point Processes*. Chapman & Hall/CRC, 1980.

[20] G. De Marco *et al.*, “Connectivity of ad hoc networks with link asymmetries induced by shadowing,” *IEEE Communications Letters*, vol. 11, no. 6, pp. 495–497, Jun. 2007.

[21] M. Desai and D. Manjunath, “On the connectivity in finite ad hoc networks,” *IEEE Communications Letters*, vol. 6, no. 10, pp. 437–439, Oct. 2002.

[22] R. Diestel, *Graph Theory*, 3rd ed. New York, NY, USA: Springer, 2005.

[23] O. Dousse *et al.*, “Impact of interferences on connectivity in ad hoc networks,” *IEEE/ACM Transactions on Networking*, vol. 13, no. 2, pp. 425–436, Apr. 2005.

[24] ——, “A case for partial connectivity in large wireless multi-hop networks,” in *Proc. UCSD Information Theory and Applications Workshop*, San Diego, CA, Feb. 6–10, 2006.

[25] ——, “Connectivity in ad-hoc and hybrid networks,” in *Proc. IEEE International Conference on Computer Communications (INFOCOM’02)*, vol. 2, New York, NY, USA, Jun. 23–27, 2002, pp. 1079–1088.

[26] D. Estrin *et al.*, “Next century challenges: Scalable coordination in sensor networks,” in *Proc. of the 5th annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom’99)*, Seattle, WA, USA, Aug. 15–19, 1999, pp. 263–270.

[27] F. Fabbri and R. Verdone, “A statistical model for the connectivity of nodes in a multi-sink wireless sensor network over a bounded region,” in *Proc. of the 14th European Wireless Conference (EW’08)*, Prague, Czech Republic, Jun. 22–25, 2008, pp. 1–6.

[28] C. R. Farrar *et al.*, “Sensor network paradigms for structural health monitoring,” *Struct. Control Health Monit.*, vol. 13, no. 1, pp. 210–225, Jan./Feb. 2006.

[29] W. Feller, *An introduction to probability theory and its applications. Volume I*, 3rd ed. New York, NY, USA: John Wiley & Sons, 1970.

[30] ——, *An introduction to probability theory and its applications. Volume II*, 2nd ed. New York, NY, USA: John Wiley & Sons, 1971.

[31] R. K. Ganti and M. Haenggi, "Single-hop connectivity in interference-limited hybrid wireless networks," in *Proc. IEEE International Symposium on Information Theory (ISIT'07)*, Nice, France, Jun. 24–29, 2007, pp. 366–370.

[32] ——, "Single-hop connectivity in interference-limited hybrid wireless networks," in *Proc. IEEE International Symposium on Information Theory (ISIT'07)*, Nice, France, Jun. 24–29, 2007, pp. 366–370.

[33] A. Ghasemi and S. Nader-Esfahani, "Exact probability of connectivity in one-dimensional ad hoc wireless networks," *IEEE Communications Letters*, vol. 10, no. 4, pp. 251–253, Apr. 2006.

[34] A. Giridhar and P. R. Kumar, "Maximizing the functional lifetime of sensor networks," in *Proc. of the 4th international symposium on Information processing in sensor networks (IPSN'05)*, Los Angeles, CA, USA, Apr. 25–27, 2005, pp. 5–12.

[35] E. Godehardt and J. Jaworski, "On the connectivity of a random interval graph," *Random Structures and Algorithms*, vol. 9, no. 1–2, pp. 137–161, Aug./Sep. 1996.

[36] A. J. Goldsmith, *Wireless Communications*, 1st ed. New York, NY, USA: Cambridge University Press, 2005.

[37] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Communications Magazine*, pp. 8–27, Aug. 2002.

[38] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," *Stochastic Analysis, Control, Optimization and Applications*, pp. 547–566, 1998, a Volume in Honor of W.H. Fleming, W.M. McEneaney, G. Yin, and Q. Zhang (Eds.) Boston, 1998.

[39] ——, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[40] P. Hall, *Introduction to The Theory of Coverage Processes*, 2nd ed. New York, NY, USA: John Wiley & Sons, 1988.

[41] W. B. Heinzelman *et al.*, “An application-specific protocol architecture for wireless microsensor networks,” *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, Oct. 2002.

[42] C. Huang and Y. Tseng, “The coverage problem in a wireless sensor network,” *Mobile Networks and Applications*, vol. 10, no. 4, pp. 519–528, Aug. 2005.

[43] J. M. Kahn *et al.*, “Next century challenges: Mobile networking for “smart dust”,” in *Proc. of the 5th annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom’99)*, Seattle, WA, USA, Aug. 15–19, 1999, pp. 271–278.

[44] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes*, 2nd ed. New York, NY, USA: Academic Press, 1975.

[45] J. F. C. Kingman, *Poisson Processes*, 1st ed., L. C. G. Rogers, Ed. New York, NY, USA: Oxford University Press, 1993.

[46] B. Krishnamachari, *Networking Wireless Sensors*, 1st ed. New York, NY, USA: Cambridge University Press, 2005.

[47] S. Kumar *et al.*, “On k-coverage in a mostly sleeping sensor network,” *Wireless Networks*, vol. 14, no. 3, pp. 277–294, Jun. 2008.

[48] J. Li *et al.*, “Meeting connectivity requirements in a wireless multihop network,” *IEEE Communications Letters*, vol. 10, no. 1, pp. 19–21, Jan. 2006.

[49] X.-Y. Li *et al.*, “Coverage in wireless ad hoc sensor networks,” *IEEE Transactions on Computers*, vol. 52, no. 6, pp. 753–763, Jun. 2003.

[50] R. Liscano *et al.*, “Mobile wireless rsa overlay network as critical infrastructure for national security,” in *Proc. IEEE International Workshop on Measurement Systems for Homeland Security, Contraband Detection and Personal Safety Workshop (IMS’05)*, Mar. 2005, pp. 96–102.

[51] M. A. Lisovich *et al.*, “Reconfiguration in heterogeneous mobile

wireless sensor networks," in *Proc. IEEE International Symposium on Wireless Pervasive Computing (ISWPC'08)*, Santorini, Greece, May 7–9, 2008, pp. 349–354.

[52] J. P. Lynch *et al.*, "Power-efficient data management for a wireless structural monitoring system," in *Proc. 4th International Workshop on Structural Health Monitoring*, Stanford, CA, USA, Sep. 15–17, 2003.

[53] ——, "Validation of a large-scale wireless structural monitoring system on the geumdang bridge," in *Proc. International Conference on Structural Safety and Reliability (ICOSSAR'05)*, Rome, Italy, Jun. 2005.

[54] S. Megerian *et al.*, "Worst and best-case coverage in sensor networks," *IEEE Transactions on Mobile Computing*, vol. 4, no. 1, pp. 84–92, Jan./Feb. 2005.

[55] S. Meguerdichian *et al.*, "Coverage problems in wireless ad-hoc sensor networks," in *Proc. IEEE International Conference on Computer Communications (INFOCOM'01)*, vol. 3, 2001.

[56] D. Miorandi and E. Altman, "Coverage and connectivity of ad hoc networks in presence of channel randomness," in *Proc. IEEE International Conference on Computer Communications (INFOCOM'05)*, vol. 1, Mar. 13–17, 2005, pp. 491–502.

[57] ——, "Connectivity in one-dimensional ad hoc networks: A queueing theoretical approach," *Wireless Networks*, vol. 12, no. 5, pp. 573–587, Oct. 2006.

[58] J. Paek *et al.*, "A wireless sensor network for structural health monitoring: Performance and experience," in *Proc. IEEE 2nd workshop on Embedded Networked Sensors (EmNetS-II'05)*, May 30–31, 2005, pp. 1–10.

[59] S. Pai *et al.*, "Transactional confidentiality in sensor networks," *IEEE Security and Privacy*, vol. 6, no. 4, pp. 28–35, Jul./Aug. 2008.

[60] S. Pattem *et al.*, "The impact of spatial correlation on routing with compression in wireless sensor networks," in *Proc. of the 3rd international symposium on Information processing in sensor networks (IPSN'04)*, Berkeley, CA, USA, Apr. 26–27, 2004, pp. 28–35.

[61] M. Penrose, *Random Geometric Graphs*, 1st ed., L. C. G. Rogers, Ed. New York, NY, USA: Oxford University Press, 2003.

[62] M. Rahimi *et al.*, "Studying the feasibility of energy harvesting in a mobile sensor network," in *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on*, vol. 1, 14-19 2003, pp. 19 – 24 vol.1.

[63] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed., T. S. Rappaport, Ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.

[64] S. I. Resnick, *Adventures in Stochastic Processes*, 1st ed. Boston, MA, USA: Birkhaüser, 1992.

[65] K. Römer and F. Mattern, "The design space of wireless sensor networks," *IEEE Wireless Communications Magazine*, vol. 11, no. 6, pp. 54–61, Dec. 2004.

[66] S. M. Ross, *Stochastic processes*, 2nd ed. Chichester, England: John Wiley & Sons, 1995.

[67] S. Roundy *et al.*, "Power sources for wireless sensor networks," *Wireless Sensor Networks*, pp. 1–17, 2004.

[68] P. Santi, "Topology control in wireless ad hoc and sensor networks," *ACM Computing Surveys*, vol. 37, no. 2, pp. 164–194, Jun. 2005.

[69] A. Scaglione and S. D. Servetto, "On the interdependence of routing and data compression in Multi-Hop sensor networks," *Wireless Networks*, vol. 11, no. 1–2, pp. 149–160, Jan. 2005.

[70] G. Sharma and R. R. Mazumdar, "A case for hybrid sensor networks," *IEEE/ACM Transactions on Networking*, vol. 16, no. 5, pp. 1121–1132, Oct. 2008.

[71] A. F. Siegel and L. Holst, "Covering the circle with random arcs of random sizes," *Journal of Applied Probability*, vol. 19, no. 2, pp. 373–381, Jun. 1982.

[72] W. L. Stevens, "Solution to a geometrical problem in probability," *Ann. Eugenics*, vol. 9, pp. 315–320, 1939.

[73] D. Stoyan and S. Helga, *Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics*, 1st ed. Chichester, England: John Wiley & Sons, 1994.

[74] D. Stoyan *et al.*, *Stochastic Geometry and its Applications*, 2nd ed. Chichester, England: John Wiley & Sons, 1995.

[75] P. Stuedi *et al.*, "Connectivity in the presence of shadowing in 802.11 ad hoc networks," in *Proc. IEEE Wireless Communications and Networking Conference, (WCNC'05)*, vol. 4, Mar. 13–17, 2005, pp. 2225–2230.

[76] D. Tian and N. D. Georganas, "Connectivity maintenance and coverage preservation in wireless sensor networks," *Ad Hoc Networks*, vol. 3, pp. 744–761, 2005.

[77] P.-J. Wan and C.-W. Yi, "Coverage by randomly deployed wireless sensor," *IEEE Transactions on Information Theory*, vol. 14, no. 6, pp. 2658–2669, Jun. 2006.

[78] L. Wang and Y. Xiao, "A survey of Energy-Efficient scheduling mechanisms in sensor networks," *Mobile Networks and Applications*, vol. 11, no. 5, pp. 723–740, Oct. 2006.

[79] X. Wang *et al.*, "Integrated coverage and connectivity configuration in wireless sensor networks," in *Proc. of the 1st international conference on Embedded networked sensor systems (SenSys'03)*, Los Angeles, CA, USA, Nov. 5–7, 2003, pp. 28–39.

[80] A. Wheeler, "Commercial applications of wireless sensor networks using ZigBee," *IEEE Communications Magazine*, vol. 45, no. 4, pp. 70–77, Apr. 2007.

[81] A. Wigderson, "The complexity of graph connectivity," *Mathematical Foundations of Computer Science*, vol. 629, pp. 112–132, 1992.

[82] N. Xu *et al.*, "A wireless sensor network for structural monitoring," in *Proc. of the 2nd international conference on Embedded networked sensor systems (SenSys'04)*, Baltimore, MD, USA, Nov. 3–5, 2004, pp. 13–24.

[83] Y. Xu *et al.*, "Geography-Informed energy conservation for ad hoc

routing," in *Proc. of the 7th Annual International Conference on Mobile Computing and Networking (MobiCom'01)*, Rome, Italy, Jul. 16–21, 2001, pp. 70–84.

[84] Y. Xu and W. Lee, "Exploring spatial correlation for link quality estimation in wireless sensor networks," in *Proc. of the 4th Annual IEEE International Conference on Pervasive Computing and Communications (PerCom'06)*, Pisa, Italy, Mar. 13–17, 2006, pp. 200–211.

[85] 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, Mar. 2004.

[86] J. Yen, "Emerging technologies for homeland security," *Communications of the ACM*, vol. 47, no. 3, pp. 33–35, Mar. 2004.

[87] R. Zheng *et al.*, "Optimal block design for asynchronous wake-up schedules and its applications in multihop wireless networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 9, pp. 1228–1241, Sep. 2006.

[88] J. Zhu and S. Papavassiliou, "On the connectivity modeling and the Trade-offs between reliability and energy efficiency in large scale wireless sensor networks," in *Proc. IEEE Wireless Communications and Networking Conference, (WCNC'03)*, vol. 2, Mar. 2003, pp. 1260–1265.