

# On the Implications of the Log-normal Path Loss Model: An Efficient Method to Deploy and Move Sensor Motes

Yin Chen  
Computer Science Department  
Johns Hopkins University  
Baltimore, MD 21218, USA  
yinchen@cs.jhu.edu

Andreas Terzis  
Computer Science Department  
Johns Hopkins University  
Baltimore, MD 21218, USA  
terzis@cs.jhu.edu

## Abstract

IEEE 802.15.4 links can be classified into three distinct reception regions: connected, transitional, and disconnected. The transitional region is large in size and characterized by the existence of links with intermediate reception ratios. Our work leverages previous work on understanding the properties of wireless links in the space and time domains but differs in the sense that we seek opportunities to *actively* adjust the physical topologies of sensor networks to improve link quality. Based on an existing theoretical model supported by extensive experiments in a variety of environments, we propose an efficient mechanism to identify locations with high reception ratios in the transitional region. The proposed mechanism can be used to effectively construct long, yet high reception ratio links that are 100% longer than the size of the connected region, thereby reducing the number of relay nodes necessary to interconnect sparse sensor networks by 34%. Furthermore, this mechanism can help better position mobile sinks and guide the communication protocols for mobile sensor networks. Overall, this paper provides fresh insights into the implications of the log-normal path loss model on deploying and moving sensor motes.

## Categories and Subject Descriptors

C.2.1 [Computer-Communication Networks]: Network Architecture and Design—*Wireless communication*

## General Terms

Design, Experimentation, Performance

## Keywords

Sensor Network, Relay Node Placement

## 1 Introduction

IEEE 802.15.4 links have been well studied over the past few years. Prior research investigated the correlation between link length and packet reception ratio and ob-

served that as length increases, wireless links go through three distinct regions: *connected*, *transitional*, and *disconnected* [41, 43]. The packet reception ratio (PRR) of short links should be close to 100% with high probability; whereas the PRR of long links should be close to 0%. However, for a link of intermediate length, one cannot predict its PRR by its length alone. This unpredictability is the defining characteristic of the transitional region. Unfortunately, the transitional region is large in size, including links of various lengths.

The log-normal path loss model [28] is a popular radio propagation model that can explain the formation of those three regions [24, 42]. According to this model, the strength of a radio signal decays with some power of distance. In addition, the signal strength is also coupled with random Gaussian variations. Those random variations make the PRR unpredictable in the transitional region and complicate the analysis of wireless link behavior greatly.

Therefore, recent efforts were devoted to understanding the behavior of wireless links, in both the space and the time domains. Patwari and Agrawal [24] observed correlated shadowing and applied its effects on signal strengths to RF tomography. Srinivasan et al. [35] proposed a metric to capture correlations in PRR on different links and used this metric to predict the performance of various protocols. In the time domain, Srinivasan et al. [36] studied the temporal properties of wireless links and demonstrated the benefits of sending packets in bursts over intermediate links. These links have border-line RSSI values and are in the transitional region, thus signal strength variations of a few dB can cause dramatic fluctuations in packet reception ratios. Therefore, sending packets during the bursty windows effectively makes use of the good moments of a bad link and avoids the bad moments of a good link.

Rather than looking at signal strength variations in the time domain, in this paper we focus on variations in the space domain. Specifically, we study the implications of the log-normal path loss model when deploying or moving sensor motes. The key difference from previous work on wireless link behaviors is that we seek opportunities to *actively* adjust the physical topologies of sensor networks, whereas previous efforts *passively* accept the existing physical topology and strive to make the best of it.

We use a running example throughout the paper, in which we consider a pair of motes placed at two separate locations with one mote sending packets to the other. Moreover, the

received signal strength at the receiver is close to the noise floor and therefore the packet reception ratio is low. Given this scenario, we would like to answer the following question: *how can we adjust the receiver's location to increase the packet reception ratio?*

Let us suppose that the log-normal path loss model holds. Accordingly, the observed signal strength at the receiver is a Gaussian random variable. Note that the stochasticity is over the space but not the time domain. Therefore when we move the receiver to a different location, its received signal strength is a new realization of the Gaussian distribution and more importantly is independent of the signal strength at the previous location. As a result, by placing the receiver at several nearby locations we can instantiate multiple realizations of the Gaussian distribution and select the location that measures the highest reception ratio. The independence of those realizations implies that finding the location whose reception ratio is above a certain threshold can be modeled as a sequence of Bernoulli trials. Therefore, the number of attempts required to find a location with high reception ratio is geometrically distributed. For example, if 25% of the vicinity has high reception ratio, then the expected number of locations to try is four. Moreover, since geometric distributions do not exhibit long tails, the actual realization, i.e., the number of trials to find a good location, will not deviate significantly from the expected value.

This simple *Bernoulli trial* approach has a profound impact. Essentially, it provides a theoretical foundation for the trial-and-error approach of placing motes: if there is a certain percentage of locations with high reception ratio in the vicinity, then trying several random locations is the efficient way to discover a good one. This technique is also useful for positioning mobile sinks [37] and the nodes of robotic sensor networks [38, 40]. At the same time, it also provides a simple way to find good receiver locations in a link's transitional region and therefore assist in the proactive construction of long, yet reliable links. This can be useful when deploying sensor motes, and especially so when deploying relay motes to interconnect the sensing motes [5].

Nevertheless, the log-normal path loss model deviates from reality because it does not consider spatial correlations [24]. Indeed, we empirically verify the existence of correlation. However, we also find that the correlation decreases as the spatial separation grows and can be ignored for practical purposes when the separation is larger than one meter for the variety of environments that we have tested.

Our experimental results indicate that the number of locations one needs to test in order to achieve high packet reception ratio is indeed geometrically distributed and that this number is usually quite small. As a rule of thumb, when placing the receiver at a distance that is twice the length of the connected region, one can find a good location in fewer than five trials.

**Contributions.** This paper makes four research contributions: **(1)** We realize and investigate the implications of the log-normal path loss model on placing and moving wireless sensor motes. **(2)** We present experimental results based on extensive measurements from near-ground 802.15.4 networks deployed in a variety of environments, ranging from

flat parking lots with few obstacles to lush forests, to validate the applicability of the proposed Bernoulli trial mechanism in realistic deployment scenarios. **(3)** We show how this Bernoulli trial mechanism, coupled with a relay placement algorithm can significantly reduce the number of required relay motes. This mechanism can also reduce the radios' transmission power. **(4)** We identify that the Bernoulli trial approach is an efficient and effective tool for positioning mobile sinks and the nodes of robotic sensor networks.

This paper has six additional sections. The section that follows introduces background material on radio signal propagation models and their application to packet loss estimation. Section 3 analyzes the proposed methodology for determining high reception ratio locations. We present our evaluation in Section 4 and describe the applications of the proposed Bernoulli trial method in Section 5. We review related work in Section 6 and close in Section 7 with a summary.

## 2 Background

This section summarizes existing probabilistic models that describe signal strength and subsequently packet loss rates as a function of sender-receiver distance.

### 2.1 Radio Channel Model

Received signal strength is commonly estimated using two complementary families of models: *large scale* models that predict signal attenuation over long distances (i.e., distances that are multiples of the signal's wavelength ( $\lambda$ )) and *small scale* models that estimate signal loss over small distances and small time intervals<sup>1</sup>.

A commonly used large scale model is the log-distance path loss model with log-normal shadowing (also known as log-normal path loss model [28]). According to this model, the received signal strength  $P_r(d)$  (in dBm) at a given distance  $d$  from the transmitter is given by:

$$P_r(d)[dBm] = P_t[dBm] - \overline{PL}(d_0) - 10n \log\left(\frac{d}{d_0}\right) - X_\sigma \quad (1)$$

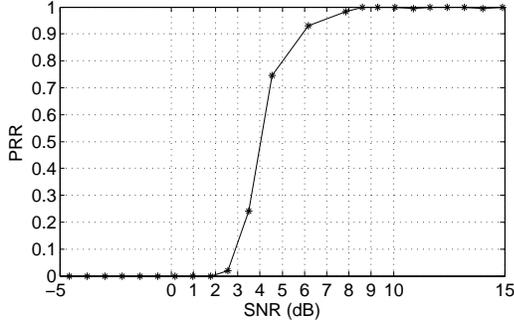
where  $P_t$  is the power of the transmitted signal,  $\overline{PL}(d_0)$  is the measured path loss at reference distance  $d_0$ ,  $n$  is the path-loss exponent, and  $X_\sigma \sim N(0, \sigma)$  is a normal random variable (in dB). The term  $X_\sigma$  models the path loss variation across all locations at distance  $d$  from the source due to *shadowing*, a term that encompasses signal strength variations due to artifacts in the environment (i.e., occlusions, reflections, etc.). Accordingly, received signal strengths at locations that are of equal distance from the transmitter are considered i.i.d. normal random variables.

While (1) accounts for signal variations over large scales, the received signal strength can vary considerably over small distances (in the order of  $\lambda$ ) and small time scales, due to *multipath fading* [28]. As a result, packet loss can exhibit wide variations even when  $d$  changes by as little as a few centimeters in the case of 802.15.4 radios.

### 2.2 Packet Reception Ratio

The received power  $P_r(d)$  can be used to estimate a link's packet reception ratio (PRR). Specifically, the PRR depends

<sup>1</sup>If the transceivers and the environment are static then so-called small-scale *fading* is a purely spatial phenomenon.



**Figure 1. Measured packet reception ratio (PRR) as a function of Signal to Noise Ratio (SNR) for the CC2420 802.15.4 radio. Packet size is set to 114 bytes.**

on the radio’s modulation scheme, the packet size, and the signal to noise ratio (SNR)  $s = P_r(d) - P_{noise}$ .  $P_{noise}$  models the power of the ambient noise, usually as a Gaussian random variable with zero mean and deviation  $\sigma_n$  [28].

We experimentally derive the mapping between PRR and SNR in the case of the CC2420 802.15.4 radio [39] that we use and Figure 1 illustrates the result. For this experiment we place the receiver at a fixed distance from the transmitter and vary the power of the noise emitted by a radio frequency signal generator, using the configuration used in [8]. The noise signal has a flat power spectral density within the operating frequency of the CC2420 radio and the ambient noise is much lower than the generated noise signal. Note that if  $\text{SNR} \geq 8$  dB then  $\text{PRR} \approx 1$ . We will use this figure as the reference SNR-PRR curve throughout Section 3. We set the packet size to 114 bytes because it is close to the maximum packet size for 802.15.4 radios (127 bytes) and a link that yields high PRR under this packet size will certainly have high PRR for smaller packet sizes.

We say a PRR is high if  $\text{PRR} \geq p_{high}$ , a predefined threshold. In the analyses and experiments that follow, we set  $p_{high} = 0.85$ .

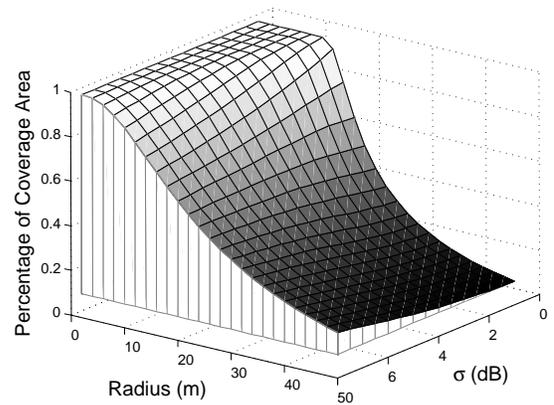
### 3 Bernoulli Trials and Good Locations

Next, we present a detailed analysis for the proposed Bernoulli trial method for the two-mote example from Section 1 and then derive the probabilities to discover good locations. We term receiver location as good if the  $\text{PRR} \geq p_{high}$ .

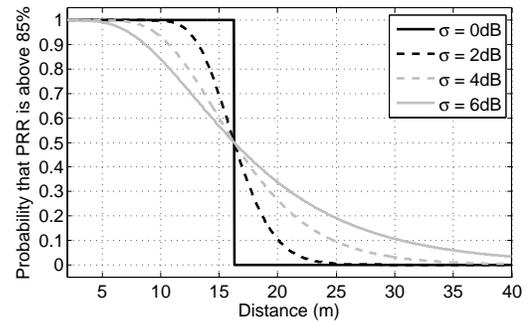
#### 3.1 Coverage Percentage and Three Regions

In what follows, we first explore the coverage percentage of the area surrounding a sender and then show the probability for a link to have  $\text{PRR} \geq p_{high}$  as a function of inter-node distance. This probability will then be used to determine the lengths for the three regions. Here coverage percentage is defined as the portion of an area that has high PRR and we say that any point in this portion is a *covered location*, i.e., a good location.

One can use (1) to calculate the percentage  $U(\gamma_{th}, R_i)$  of the area defined by a disk of radius  $R_i$  centered at the sender that has  $\text{SNR } s \geq \gamma_{th}$  or equivalently  $\text{PRR} \geq p_{high}$ . The cov-



**Figure 2. The percentage  $U$  of a circular area that has  $\text{PRR} \geq p_{high} = 0.85$ , as a function of radius  $R_i$  and path loss variation  $\sigma$ . The parameters for the log-normal path loss model were experimentally derived using data from an outdoors experiment using 802.15.4 radios.**



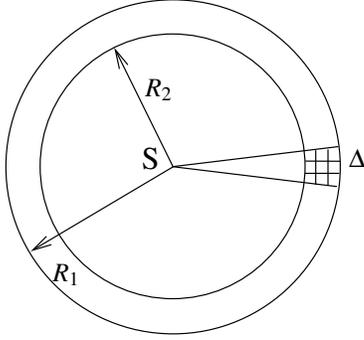
**Figure 3. The probability for a link to have  $\text{PRR} \geq p_{high} = 0.85$ , as a function of transmitter-receiver distance under various log-normal variations  $\sigma$ . The model parameters are the same as those used in Figure 2.**

erage percentage  $U(\gamma_{th}, R_i)$  is given by:

$$U(\gamma_{th}, R_i) = \frac{1}{2} \left[ 1 - \text{erf}(P_i) + \exp\left(\frac{1 - 2P_i Q_i}{Q_i^2}\right) \left[ 1 - \text{erf}\left(\frac{1 - P_i Q_i}{Q_i}\right) \right] \right] \quad (2)$$

where  $P_i = [\gamma_{th} - P_i + \overline{PL}(d_0) + 10n \log(\frac{R_i}{d_0})] / (\sigma\sqrt{2})$ ,  $Q_i = (10n \log e) / (\sigma\sqrt{2})$ , and  $\text{erf}(\cdot)$  is the error function [28].

Figure 2 draws  $U(\gamma_{th}, R_i)$  as a function of  $R_i$  and the path loss variation  $\sigma$ . As expected,  $U$  shrinks as  $R$  grows. The effect of  $\sigma$  on  $U$  is more interesting. Over short distances, higher values of  $\sigma$  result in lower coverage  $U$ . However as  $R$  increases, increased path loss variance leads to increased coverage. This seemingly counter-intuitive relationship can be explained as follows. When the signal strength is already low (i.e., at long distances), further decreases due to shadowing will decrease PRR only slightly (see Fig.1). On the other hand, equal increases in received signal strength will push the SNR and therefore the PRR upwards.



**Figure 4.** Track defined by two concentric circles centered at  $S$  with radii  $R_1 > R_2$ . A sector within that track is highlighted with a  $3 \times 3$  grid. The dimensions of each grid element are  $\Delta \times \Delta$  meters.

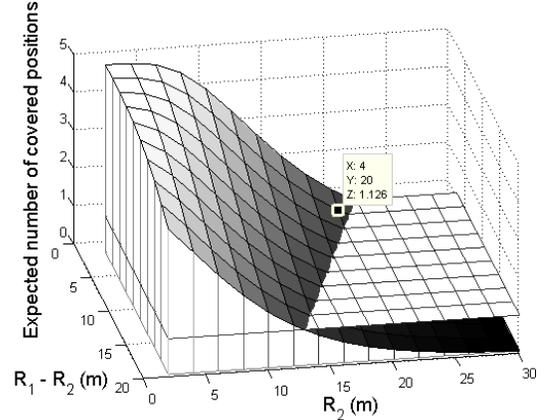
While Figure 2 shows the impact of  $\sigma$  on coverage, Figure 3 illustrates the effect of  $\sigma$  on the lengths of the three regions. Specifically, it shows the probability for a link to have  $\text{PRR} \geq p_{\text{high}}$  as a function of the sender-receiver distance, under various log-normal variations  $\sigma$ . When  $\sigma = 0$ , there is no log-normal shadowing and thus signal strength decays monotonically with distance. Therefore, the probability becomes zero as signal strength falls below  $\gamma_{\text{th}}$ , and the corresponding distance is the length of the connected region. There is no transitional region in this case. On the other hand, when  $\sigma > 0$ , determining the three regions becomes more subjective. We define the length of the connected region as the distance where the probability for a link to have  $\text{PRR} \geq p_{\text{high}}$  is equal to  $\zeta$ . For example, when  $\sigma = 4$  dB and  $\zeta = 0.95$ , the length of the connected region is  $\approx 10$  meters in Figure 3. Similarly, we define the transitional region as the range of the distance in which the probability for a link to have  $\text{PRR} \geq p_{\text{high}}$  is less than  $\zeta$  but higher than  $1 - \zeta$ . When  $\sigma = 4$  dB and  $\zeta = 0.95$ , the transitional region approximately spans from 10 meters to 30 meters. One can see that the transitional region is indeed large in size, and in this example, it is twice the size of the connected region.

Next, we will derive the probabilities of finding good locations. As the first step, we will use (2) to estimate the coverage percentage for the area between two concentric circles with radii  $R_1$  and  $R_2$  ( $R_1 > R_2$ ) centered at the sender. We term the area between the two circles a *track*. It can be shown that the coverage percentage in this case is equal to:

$$U(\gamma_{\text{th}}, R_1, R_2) = \frac{R_1^2 U(\gamma_{\text{th}}, R_1) - R_2^2 U(\gamma_{\text{th}}, R_2)}{R_1^2 - R_2^2}. \quad (3)$$

Assuming homogeneity, (3) also estimates the coverage of any sector within the track defined by  $R_1$  and  $R_2$ . Figure 4 presents such a track defined by two concentric circles centered at  $S$  as well as one sector within that track.

Let us consider the scenario in which we place the receiver mote within a track sector of the area defined by the two concentric circles. The probability that a certain position is covered (i.e.,  $\text{PRR} \geq p_{\text{high}}$ ) is approximated  $p = U(\gamma_{\text{th}}, R_1, R_2)$ . Then, assuming independence, the probability that we will find at least one position which is covered in



**Figure 5.** Expected number of covered positions ( $\text{PRR} \geq p_{\text{high}}$ ) encountered by  $N = 5$  random trials in the area defined by concentric circles  $R_1$  and  $R_2$ . The effective communication range (20 m) increases by  $\sim 100\%$  compared to the range defined by the naive approach (10 m). We use the same  $\gamma_{\text{th}}$  and  $p_{\text{high}}$  values as in Figure 2.

$N$  trials is  $1 - (1 - p)^N$ , while the expected number of positions found is  $N \cdot p$ .

Figure 5 plots the expected number of good locations as a function of  $R_2$  and  $R_1 - R_2$ , using the parameters from Figure 2 and  $N = 5$ . It is evident that even with few trials, it is still likely to find a covered location when  $(R_1, R_2) = (24, 20)$  m). Therefore, one can actually place the receiver at a distance of 20 to 24 meters away from the sender, while ensuring that  $s \geq \gamma_{\text{th}}$  (and thus  $\text{PRR} \geq p_{\text{high}}$ ). Doing so increases the effective communication range by  $\sim 100\%$ , compared to the length of the connected region determined by the log-normal path loss model (10 m). The trade-off is the need to test several locations before discovering a covered location. Nevertheless, we can control the number of trials by selecting an appropriate value for  $p$ . In turn, selecting the values of  $R_1$  and  $R_2$  determines the value of  $p$ , and also decides the sender-receiver distance. Finally, we note that trials are still necessary even if one places the receiver at a distance equal to the length of the connected region, because the probability of finding a link there with  $\text{PRR} \geq p_{\text{high}}$  is  $\zeta < 1$ .

### 3.2 Grid Sampling

The analysis so far assumes that signal strengths in nearby locations are independent random variables, controlled only by the distances between these points and the transmitter. This assumption however is not true, as signal strengths in nearby locations are correlated due to the common set of scatterers and occlusions [24]. Nonetheless, Puccinelli and Haenggi observed that displacements of  $\lambda/2$  (i.e.,  $\sim 5$  cm for 802.15.4 radios) are sufficient to guarantee the independence of multipath fading even in the presence of shared occlusions for 802.15.4 radios [27]. Our experimental results (see Section 4.3) verify that signal strengths display strong spatial correlations and therefore in some cases displacements of  $\lambda/2$  are adequate to ensure independence, but not in all cases. Nevertheless, for all the environments that we tested, a *one meter* displacement removes virtually all correlations.

Config.	Description
1	Line topology, Elevation 0 cm $\{P, L\}$
2	Line topology, Elevation 15.1 cm $\{P, L, O\}$
3	Line topology, Elevation 31.2 cm $\{P\}$
4	Tee topology, Elevation 0 cm $\{P, H, O\}$
5	Tee topology, Elevation 15.1 cm $\{P, H\}$
6	Real-life deployment $\{FA, FB\}$

**Table 1. Test sites and node configurations.**

Let us denote by  $\Delta$  the displacement that is necessary to remove any correlations. Accordingly, rather than sampling random locations within the track, one needs to select locations that are separated by multiples of  $\Delta$ . For this reason, we embed a rectangular grid with dimensions of  $n \times \Delta$  by  $m \times \Delta$  within the track, centered at the receiver’s original location, and use the grid to sample locations (see Fig.4). Then, given coverage probability  $p$ , the probability of finding at least one covered point is  $1 - (1 - p)^{n \cdot m}$ . Note that the separation of  $\Delta$  makes this grid sampling a Bernoulli trial, and therefore we call it the *Bernoulli Grid* method hereafter.

It is easy to see that even for small values of  $p$  the probability of detecting at least one good location increases rapidly with the number of trials. As we will show in Section 4.6, a small number of trials is enough in practice. Finally, the total number of trials necessary to identify a covered location are geometrically distributed with expectation equal to  $1/p$ , which we will experimentally verify in Section 4.6.

Another limitation of the log-normal path loss model is that it does not consider any temporal RF channel variations. On the other hand, multiple studies have experimentally observed and modeled the temporal variations of low-power wireless links [7, 34, 36]. Nevertheless, signal strength variations in the time domain are mostly due to time-varying levels of interference and movements in the surrounding environment. In other words, unless the environment permanently changes after the receiver has been placed, the selected locations selected will have high PRR on average. The results from Section 4.4 validate this hypothesis.

## 4 Evaluation

In this section we will first verify the log-normal path loss model for a variety of environments. Next, we investigate the spatial correlations for received signal strength as well as packet reception ratio, and find the sufficient separation between two locations that will ensure the correlation is sufficiently low. Then, we show that the long links constructed through the Bernoulli Grid method are stable over time and symmetric. Finally, we show the average number of trials to get a good location when connecting to one or two motes.

### 4.1 Methodology

We conducted tests in five increasingly complex RF environments: an outdoor parking lot spanning approximately 600 m<sup>2</sup>, an open lawn area with a higher diffraction index than the parking lot, a building hallway with line-of-sight transmissions, an indoors testbed deployed over multiple offices, and finally sensor networks deployed in two forests.

All experiments use TelosB [26] motes equipped with IEEE 802.15.4 compliant TI CC2420 radios [39]. We use two network topologies to measure the log-normal path loss parameters for the parking lot, hallway and indoors testbed environments. The first is the *line topology* in which every mote takes its turn as the transmitter. The distance between adjacent nodes in this case is 100 to 280 cm. The other is the *tee topology* in which all the receivers form a straight line, 15 to 100 cm apart from each other, while the transmitter moves away from the receivers on a trajectory that is perpendicular to the line of the receivers.

We automated the data collection process in the second topology by connecting the transmitter to an Ebox-3854 Linux PC mounted on top of an iRobot Create robot [15]. At each measurement step, the robot moves by  $D_r = 5$  cm and pauses to allow the transmitter to send a batch of packets. The benefit of the second topology is that it enables us to easily collect a large number of measurements distributed over space. Only the line topology was used in the lawn environment. All experiments use 114-byte packets and each transmitter sends at least 500 packets per batch, with the inter-packet-interval (IPI) set to 500 milliseconds. Here we set the IPI to 500 ms such that we will not be measuring the PRR of bursty windows [36]. Upon receiving a packet, each receiver records the corresponding source ID, RSSI, LQI, sequence number, and batch number. The receiver also samples the ambient noise level immediately after each packet reception.

Table 1 summarizes the different experiment configurations we used. The elevation listed in this table corresponds to the motes’ vertical distance from the ground. All motes were placed at the same height for the line topologies. On the other hand, only the receivers are elevated to the listed height for the tee topologies. The transmitter is carried by the robot and its height is fixed at 12 cm. Last, symbols  $P, L, H, O$  and  $F$  correspond to the parking lot, lawn, hallway, office, and the forest environments respectively. For the two forests, we used the measurements collected by the motes deployed there for the purpose of environmental monitoring.

### 4.2 Log-Normal Path Loss

The first step in validating the probabilistic model proposed in Section 3 is to verify the log-normal path loss model in the environments summarized by Table 1. According to this model, the average received signal strength  $\mu(d)$  decreases with a power of the distance, while the received signal strength at distance  $d$  is a Gaussian random variable,  $P_r(d) \sim N(\mu(d), \sigma)$ , due to the shadowing effects. From (1),  $\mu(d)$  is:

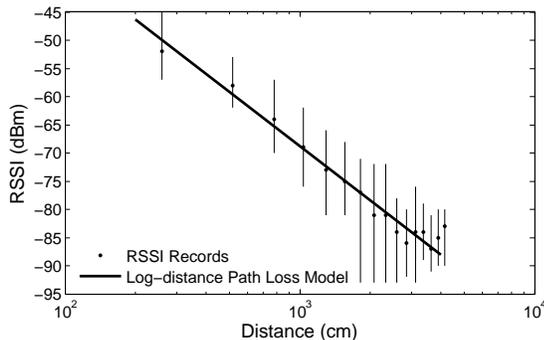
$$\begin{aligned} \mu(d) &= P_t - \overline{PL}(d_0) - 10n \log\left(\frac{d}{d_0}\right) \\ &= P_r(d_0) - 10n \log\left(\frac{d}{d_0}\right) \end{aligned} \quad (4)$$

Therefore, the average RSSI for a transmitter-receiver pair separated by distance  $d$  can be expressed using constant  $d_0$  and two parameters  $(P_r(d_0), n)$  that can be derived through linear regression over the collected measurements.

Table 2 presents the results of such linear regressions across the different environments. Letters denote the type of environments while the number that follows these letters specifies the topology configuration as listed in Table 1.

$n$	$P_r(2m)$	Configuration
2.05	-53.5 dBm	H4
1.74	-51.6 dBm	H5
1.68	-61.8 dBm	H5, Power 18
3.75	-62.9 dBm	L1
3.45	-49.9 dBm	L2
3.03	-76.3 dBm	P1
2.78	-56.2 dBm	P2
3.17	-45.2 dBm	P3
1.97	-74.9 dBm	O4, Power 3
1.97	-61.4 dBm	FA6
2.19	-53.4 dBm	FB6

**Table 2. Large scale fading parameters derived from the different environments. Transmission power levels other than the default (31) are explicitly indicated.**

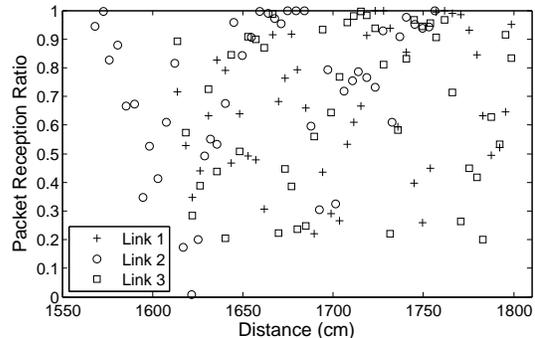


**Figure 6. Average and spread of measured RSSI values for the parking lot experiment. The line corresponds to the linear regression fit for the log-normal path loss model. Both axes are in logarithmic scale.**

It is evident that different environments produce different  $n$  and  $P_r(d_0)$ . Moreover, increasing the nodes' elevation in the same environment leads to longer communication ranges, as indicated by the higher  $P_r$  values. Hallway experiments tend to have smaller  $n$  values, a finding shared by previous measurement studies [28]. Interestingly,  $n$  values from the two forests are smaller than the path loss exponents at the parking lot and lawn sites and closer to those in the hallway and office sites.

Figure 6 plots  $P_r(d)$  as a function of distance  $d$ , using data collected from a parking lot experiment. One can see that experimental data match well the log-normal path loss model. Also evident is the *long tail* of RSSI values as  $d$  increases, where the sample average of collected RSSIs is higher than the predicted mean signal strength  $\mu(d)$ . One possible reason for this discrepancy is that RSSI records are available only for successfully received packets. For this reason, RSSI measurements from links with low reception ratios are biased towards higher values.

Next, we verify that RSSI variations indeed follow the



**Figure 8. Measured Packet Reception Ratios within the radio's transitional region as a function of transmitter-receiver distance. PRR samples from three receivers are collected as the transmitter moves at 5 cm intervals in the hallway environment. Tx power is set to 3.**

log-normal distribution used by the model. To do so, we use (4) to calculate  $\mu(d)$  for each experiment, using the parameters listed in Table 2. Then, we subtract the predicted mean RSSI  $\mu(d)$  from the corresponding raw RSSI samples. The resulting *residual* RSSI should then be samples of the random variable  $X_\sigma$  in (1).

Figure 7(a)-(c) show the Quantile-Quantile plot of residual RSSI in three environments versus the standard Gaussian distribution. Figure 7(d) is the aggregate curve for all the configurations listed in Table 2. One can see that the residual RSSI matches well the Gaussian distribution, although there are some discrepancies in the tails in some environments.

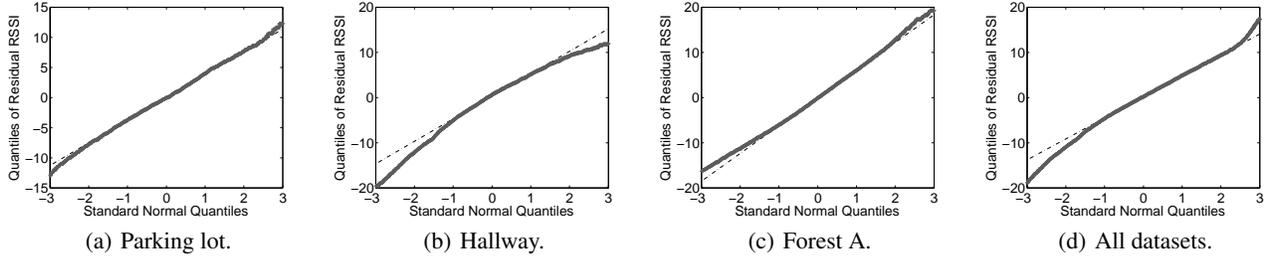
### 4.3 Spatial Correlation and Coverage

The results from the previous section confirm that the log-normal path loss model agrees with the experimental data gathered across a variety of environments. Furthermore, the variations in RSSI are indeed normally distributed. However, they are not necessarily independent. In this section, we will study the correlations among those variations. Specifically, we start by looking at the PRR variations with small changes in location.

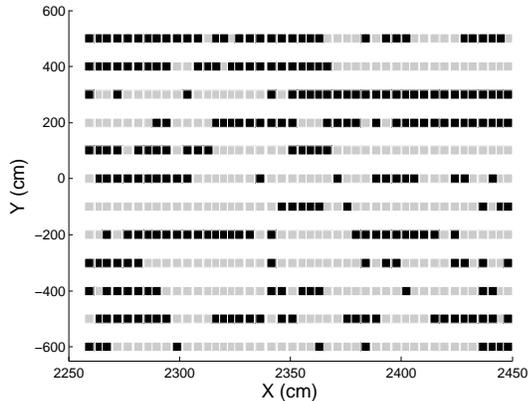
Figure 8 shows the packet reception ratios as the distance increases at increments of five centimeters. One can see that the PRR can fluctuate significantly even within a short distance. This variation is due to small-scale fading and is the defining signature of the transitional region.

Furthermore, Figure 9 presents this small-scale spatial PRR variation over a two-dimensional grid, using data collected from one parking lot experiment. The gray cells in this graph correspond to locations with  $\text{PRR} \geq 85\%$  (covered or good locations), whereas the black cells have  $\text{PRR} < 85\%$  (coverage holes)<sup>2</sup>. The interlacing of good locations and coverage holes in the transitional region agrees with the model in Section 3 and signals good news. The coverage holes are not clustered and therefore doing grid sampling in the vicinity of a coverage hole is likely to discover a good

<sup>2</sup>We observe that the low PRR measurements in Figure 9 are due to low received signal strength, rather than external noise or interference.



**Figure 7. Q-Q plot of residual RSSI versus standard normal distribution for three environments, and the aggregate curve for all the datasets listed in Table 2. In all cases, residual RSSI is well modeled by the Gaussian distribution.**



**Figure 9. Measured PRR in a parking lot environment. Black cells correspond to coverage holes, while gray cells designate good locations (i.e.,  $\text{PRR} \geq 85\%$ ). The X- and Y-axis values correspond to relative distances to the transmitter on a 2D plane.**

location. For example, in the mobile sink or robotic sensor networks scenario, a mobile mote can simply move to a nearby location when stepping in a coverage hole. Moreover, from those two figures one can see that it would be overly conservative and unnecessary to completely avoid the transitional region due to its coverage holes. After all, it is not that difficult to find a good location. Instead, we can use the variation to identify good locations and thereby construct long (distance-wise) and good (PRR) links, which is useful when deploying relay nodes (cf. Section 5).

While the log-normal path loss model considers the PRRs at different positions to be independent, Figure 9 suggests that this model is not valid when the distance between the positions is small. These dependencies were also identified by Patwari and Agrawal, who proposed a correlation model for RSSI values at nearby positions [24]. In their model the correlation exponentially decays with the distance between the two positions. Patwari and Agrawal also stated that the correlation depends on the characteristics of the environment.

We use measurements collected from different environments to characterize the correlation between the received signal strengths at two nearby locations. We do so, to estimate the sufficient distance  $\Delta$  that will guarantee statistical independence for the corresponding PRRs. This  $\Delta$  can be used in the Bernoulli Grid method presented in Section 3.2.

Specifically, we calculate the correlation coefficients for raw RSSI, PRR, and residual RSSI as a function of the distance between the two measurement locations. We calculate residual RSSI by subtracting  $\mu(d)$  from raw RSSI, in order to eliminate the correlations introduced by  $\mu(d)$ . Figures 10(a), 10(b), and 10(c) show these correlation coefficients. Figures 10(a) and 10(b) correspond to two experiments in different areas of a large parking lot. One can see that correlations decay as distance increases, reducing from 1.0 to  $\sim 0.2$ . In Figure 10(c), the correlation for PRR decays considerably at a few centimeters, suggesting  $\lambda/2 = 6\text{cm}$  is a sufficient  $\Delta$  for that environment. However, Figure 10(b) shows that it can take up to 1 meter for the correlation to decrease. Given this last result, we set  $\Delta = 1\text{m}$  for the rest of this discussion<sup>3</sup>.

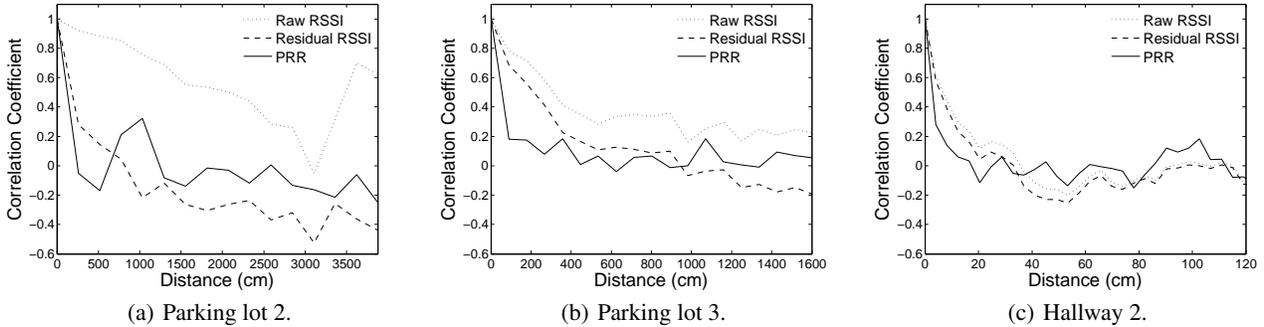
We note that Puccinelli and Haenggi observe that,  $\Delta \geq \lambda/2$  is sufficient to make the multipath fading at two positions independent [27]. The implicit requirement for this to hold is that radio waves arrive from all angles with equal probability, in which case the correlation coefficient of the fading signal can analytically be shown to be approximating zero [23]. However, in real life it is not always true that multipath signals arrive from all angles equally and therefore larger distances are necessary to remove the correlation.

#### 4.4 Temporal Variation

One implicit assumption of the proposed Bernoulli Grid method is that PRR and RSSI values observed in a relatively short period of time are accurate long term predictions of these values. On the other hand, temporal variations are often observed in 802.15.4 links [34, 36], seemingly contradicting the above assumption. Note that we have set the inter-packet-interval to 500 ms according to [36] to reduce the impact of bursty links. In this section, we first show that measuring PRR over a few minutes can predict PRR values over the following hour. Then we present RSSI measurements collected from the two forest deployments across several months which show that RSSI measurements are consistent even over very long time periods.

We start with the relationships between short-term and medium-term PRRs in four experiments, two indoor (hallway and the indoors testbed) and two at a parking lot. In each experiment, one mote is selected as the common transmitter and broadcasts packets to other motes. The transmitter is placed at various locations and broadcasts 10,000 packets with an inter-packet interval of 500 ms at each location. We

<sup>3</sup>While not shown here, one meter is also sufficient for other environments that we have tested.



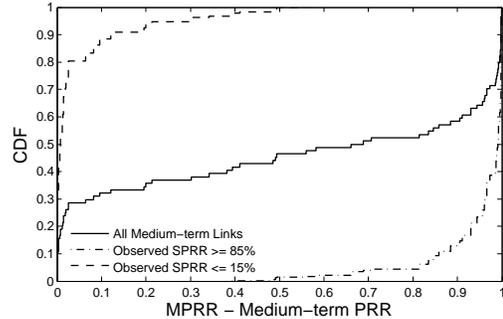
**Figure 10. Correlation coefficients between the RSSI and PRR values at two nearby locations for different environments.**

test a total of 84 links, all of which are in the transitional region, i.e., are longer than the length of the connected region as defined in Section 3.1.

We define the PRR over all the 10,000 packets as the medium-term PRR (MPRR). Furthermore, we divide the 10,000 packets into 20 non-overlapping blocks, each containing 500 packets. We define the PRR for each block as the short-term PRR (SPRR). Therefore, MPRR is the average over 20 SPRRs and we can calculate the standard deviation for the corresponding 20 SPRRs. The median SPRR standard deviation is 2%, while for 90% of all the link MPRRs the SPRR standard deviation is less than 20%. Figure 12 shows the cumulative distribution for MPRR conditioned on the observed SPRR. One can see that when  $\text{SPRR} \geq 85\%$ , the probability of  $\text{MPRR} \geq 85\%$  is  $\approx 90\%$ . These results suggest that selecting a good location based on a short duration PRR test offers a reliable prediction of the link’s quality at least for the next hour. Finally, we have verified that temporal PRR variations in our experiments were due to fluctuations in RSSI rather than ambient noise or interference. Such temporal RSSI variations in indoor environments are considerably larger during the day time when people are moving around the building.

While the previous results suggest that PRR remains fairly consistent within one hour, what happens over longer timescales? Note that for mobile sinks and robotic sensor networks, medium-term stability is more than sufficient because nodes are mobile and thus can easily relocate. Therefore, long-term stability is more relevant to deploying stationary sensor nodes and relay nodes. To study the long-term stability, we collect RSSI measurements from *FA* and *FB* from a total of 120 motes over half a year.

Specifically, for each link in the transitional region, we calculate the standard deviation of RSSI measurements across the six months. The average standard deviations across all links were 3.4 dB for *FA* and 3.7 dB for *FB*. We also collected RSSI measurements at 18 locations within *FA* for a few minutes and calculated the standard deviation of RSSI measurements for each location. The average standard deviation over these values was 2.1 dB. The fact that the long term standard deviation is similar to the short term standard deviation indicates that, for each link, short term RSSI and long term RSSI are not significantly different. Moreover, Figure 11(a) shows the daily residual RSSI pattern. One can



**Figure 12. CDF of MPRR under various SPRRs. Links with high ( $\geq 85\%$ ) SPRR are very likely to also have high MPRR.**

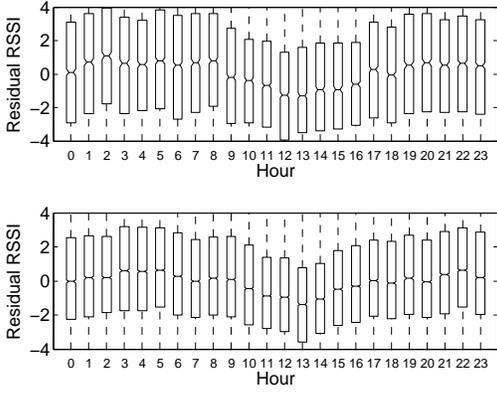
see that there is considerable variation of RSSI within a day, which also contributes to the long term standard deviation. Boano et al. [4] observe that RSSI values become smaller as the temperature rises, which explains the hourly fluctuations shown in Figure 11(a). On the other hand, as Figure 11(b) indicates, the monthly distribution is more stable.

We note that temporal properties of wireless links depend on the characteristics of the environments. The results in Figure 11(b) suggest that for sensor networks deployed in forests, link qualities are likely to be stable over time. For more volatile environments, however, the results in Section 4.3 can be used to guide the placement of sensor nodes (e.g., separated by at least 1 meter) such that correlated link failures are less likely to occur.

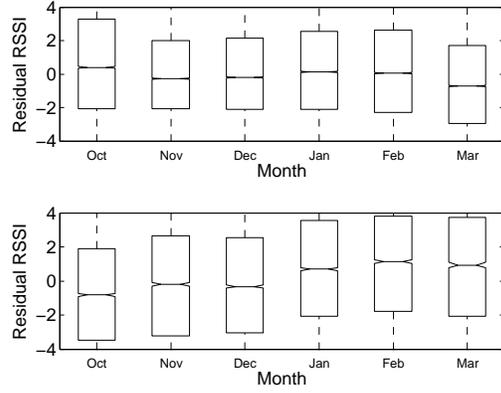
#### 4.5 Link Asymmetry

We conduct four experiments to explore whether the links selected by the proposed mechanism are asymmetric. One of the experiment is indoors (*H*), while the rest are outdoors (*P*). We test a total number of 183 links across all cases. For all experiments, we measure PRRs in both directions and investigate whether we can use the PRR from one direction to predict the PRR for the reverse direction of the same link.

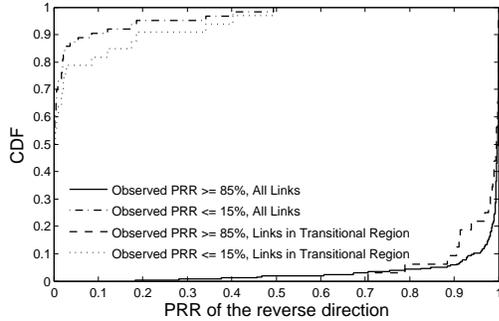
Figure 13 shows the distribution of the packet reception ratio for the reverse link direction, given that the observed link direction has  $\text{PRR} \geq 85\%$ , or  $\text{PRR} \leq 15\%$ . In each case, two curves are shown. The first curve considers all the links, including the ones in the connected region, whereas the second curve shows only the links within the transitional region.



(a) 24 hours.



(b) 6 months.

**Figure 11. Box plots of residual RSSI values over 24 hours and 6 months for the two forest sites.****Figure 13. CDF of the PRR of the reverse link direction when the forward link direction has PRR  $\geq 85\%$  or  $\leq 15\%$ . High-quality links are symmetric.**

One can see that links with high PRR tend to be symmetric: the probability that the other direction also has PRR  $\geq 85\%$  is approximately 0.95. Moreover, PRR symmetry also holds for links in the transitional region. These results agree with previous studies which showed that most asymmetries occur in links with intermediate PRRs [29, 34]. On the other hand, links with either high ( $\geq 85\%$ ) or low ( $< 15\%$ ) PRR tend to be mostly symmetric.

#### 4.6 Number of Trials

Next, we evaluate the correctness and effectiveness of applying the Bernoulli Grid method from Section 3.2 in four different environments:  $H$ ,  $O$ ,  $P$ , and  $FA$ . We are most interested in whether the number of trials to find a good location is actually geometrically distributed. Therefore, we compare the distribution of the experimentally counted number of trials necessary to find a good location to the number of trials predicted by the equations in Section 3.2.

We first measure the log-normal parameters and SNR-PRR curve for each environment. With those parameters, we can compute the length of the connected region  $l_c$  as defined in Section 3.1. Next, we fix the location of the first mote and place the second mote at a distance  $l_t$  from the fixed mote. We let  $l_t > l_c$  thus putting the mote in the transitional region. Then, we use a control mote to broadcast a beacon instruct-

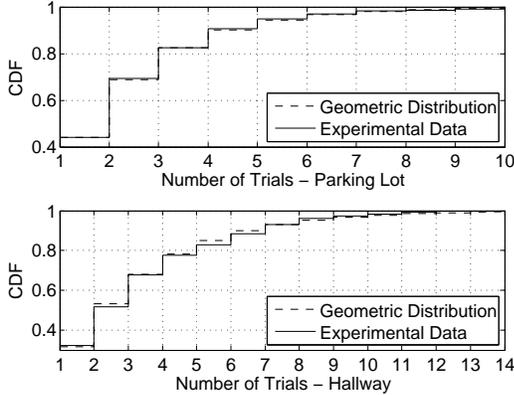
Dataset	$T$	$E(T)$	$l_c$ (cm)	$l_t$ (cm)	TX power
P4	2.26	2.16	1324	2327	23
H4	3.16	2.85	477	1550	3
H4	1.24	1.23	477	900	3
O4	2.24	1.81	300	700	3
FA6	1.67	1.81	3000	7000	31

**Table 3. Average number of trials to construct links with various lengths in four environments.  $T$  is the average number of trials, and  $E(T)$  is the expected number of trials predicted by the model.  $l_c$  and  $l_t$  are the lengths of the connected regions and the links constructed using the Bernoulli Grid method respectively.**

ing the two motes to transmit a certain number of packets, to measure the PRR of the constructed link. We move the second mote to multiple locations, iteratively constructing the Bernoulli Grid described in Section 3.1 and repeat the packet transmission process. Finally, we use different motes and put them in different areas in each environment and repeat the above process.

Through this process, we can compute the average number of trials to reach good locations for each environment, under different  $l_t/l_c$  ratios. Table 3 lists the average number of trials to get a good location in four types of environments. One can see that when  $l_t/l_c \approx 2$ , the number of trials is still quite small, which indicates that it is very easy to construct links that are twice the length of the connected region. We observe that as a general rule of thumb, when putting the receiver at a distance that is twice the length of the connected region, we can find a good location in less than five trials.

Furthermore, in the above table the average number of trials matches the expected number of trials. But we are more interested in the distribution of the number of trials, therefore we plot Figure 14 which compares the cumulative distributions of the number of trials in the parking lot and hallway experiments (the first two rows in Table 3) against the corresponding geometric distributions (whose parameter  $p$  is set



**Figure 14. Cumulative distribution of the number of trials in parking lot and hallway experiments and the corresponding Geometric distributions.**

to  $1/T$ ). Obviously, the number of trials is indeed geometrically distributed. This is a very important result because it is the theoretical foundation of the random trial method.

#### 4.7 Connecting Two Motes

All the results presented so far are limited to single hop links involving two motes, a sender and a receiver. When placing relay nodes, however, we may need to place one relay between two motes to connect with both. Next, we evaluate the effectiveness of the Bernoulli Grid method in connecting two motes using the same relay.

Let  $A$  and  $B$  be two fixed sensing motes that we need to connect using relay nodes. Furthermore, let the distance between mote  $A$  and  $B$  be  $l > l_c$ . In this case, a simple approach would be to let the distance between two adjacent motes be  $\leq l_c$  and thus would need to place  $\lceil l/l_c \rceil - 1$  relay nodes between  $A$  and  $B$ . For example, when  $l = 4l_c$ , three relay nodes are necessary. In contrast, using the Bernoulli Grid method one can let the inter-mote distance be  $2l_c$  and thereby require a single relay, placed at the midpoint between  $A$  and  $B$ .

Table 3 lists the number of trials necessary to find a location for relay  $R$  such that link  $(A, R)$  is connected for various environments. However, in order to connect  $A$  and  $B$ , we also need to connect link  $(B, R)$ . Note that a relay location that is good for link  $(A, R)$  is not necessarily good for link  $(B, R)$ . Let  $T_{AR}$  and  $T_{BR}$  be the expected numbers of trials to find a good location for link  $(A, R)$  and  $(B, R)$  respectively, and let  $T_{ARB}$  be the expected number of trials to find a location that is good for both link  $(A, R)$  and  $(B, R)$ . Assuming that the PRRs for link  $(A, R)$  and link  $(B, R)$  are independent, we should have  $T_{ARB} = T_{AR} \times T_{BR}$ .

We perform a set of experiments in which two motes are placed at two locations separated by  $l$  meters and place a relay node at the midpoint. Following the methodology described in Section 4.6, the relay mote is also placed at a series of locations forming a grid centered at the midpoint of those two endpoints motes. For each environment several TX power levels are tested, to emulate different node separations. Table 4 summarizes the corresponding number of trials. One can see that  $T_{ARB} < T_{AR} \times T_{BR}$ , which means that link  $(A, R)$  and  $(B, R)$  are not independent. Nonetheless, the

Dataset	$T_{AR}$	$T_{BR}$	$T_{AR} \times T_{BR}$	$T_{ARB}$	TX power
H4	2.72	1.61	4.39	3.99	3
H4	1.23	1.11	1.37	1.35	4
P4	2.26	1.68	3.79	3.49	23
P4	1.80	1.51	2.73	2.70	27
P4	1.60	1.43	2.31	2.28	31

**Table 4. Average number of trials for connecting two motes  $A, B$  using relay node  $R$ .  $T_{AR}$ ,  $T_{BR}$  and  $T_{ARB}$  are the average number of trials to find a good location of  $R$  that will connect link  $(A, R)$ ,  $(B, R)$  and  $(A, R, B)$  respectively.**

correlation often reduces the number of trials needed to find a location that is good for both ends. Furthermore, in all cases the necessary numbers of trials to connect both endpoints to the relay node is small.

Another interesting observation from Table 4 is that  $T_{AR} \neq T_{BR}$  even though the relay is placed at the midpoint. One possible explanation for this disparity is the non-isotropic transmission range [44] of the CC2420 radio that we use.

#### 4.8 Lowering Transmission Power

Rather than leveraging the Bernoulli Grid method to place motes farther apart than the connected region, we can lower the motes' transmission power but keep the inter-mote distance unchanged to reduce power consumption. Figure 15(b) shows that raising the transmitter's power level to 18 increases the length of the connected region, compared to Figure 15(a). However, even with the transmission power level set to 3, one can still find good positions at 15-20 meters by grid sampling, as is shown in Figure 15(a). This simple comparison suggests that the grid sampling mechanism can match the range of the connected region of higher transmission power, while using lower transmission power. In turn, the radio's transmit current draw for the two power levels is 13.65mA and 8.5mA respectively, or a 38% reduction in radio transmission power consumption. We note that transmission power is only a part of the radio power footprint and its consumption scales with traffic loads. Nevertheless, lower transmission power implicitly reduces interference to other motes and thereby the idle listening power consumption for duty-cycled MAC protocols, especially the sender-initiated ones (e.g., LPL [25]) as shown in [9].

### 5 Applications

In this section we will present the applications of the proposed Bernoulli trial method. First, we show that integrating this method into a site planning tool can significantly reduce the number of relay nodes that we need to deploy. Second, we show that this method can be used to position mobile sinks and assist networks of mobile sensing robots.

#### 5.1 Placing Relay Nodes

Figure 16 presents an environmental monitoring network deployed in a forest in Edgewater, MD to measure soil conditions. The network consists of sensing motes deployed in multiple distinct sensor patches and a number of relay nodes deployed to connect these patches into a single network rooted at the gateway located at the top left part of the

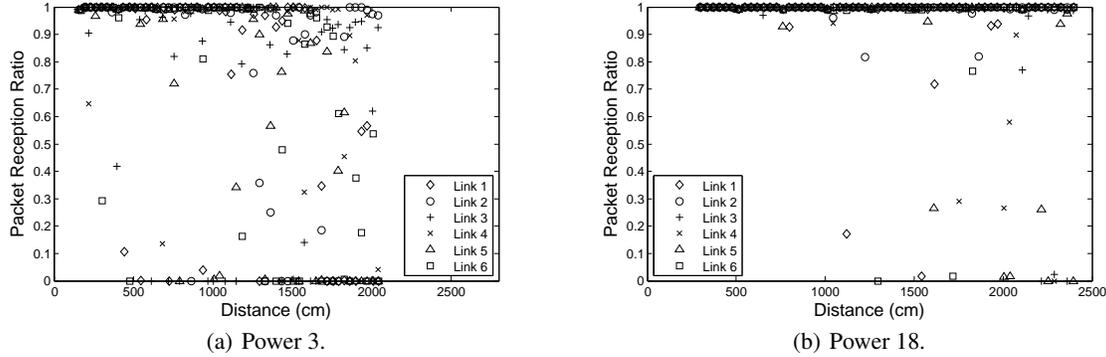


Figure 15. PRR vs. distance for six links at two transmission power levels.

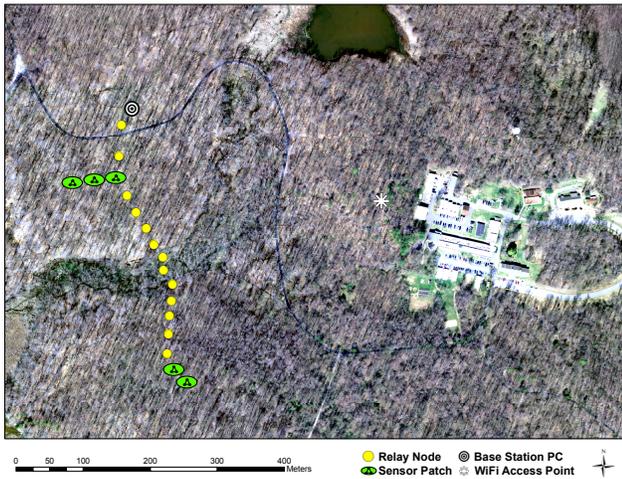


Figure 16. Map of an environmental sensing network deployment including sensors deployed in multiple network patches and a series of relay nodes connecting these patches to the network’s gateway.

map. The gateway is connected via a long-range WiFi link to an Internet-connected PC. It is not possible to connect the sensor patches using WiFi links, due to the lack of power.

This deployment provides one example of *multi-resolution sensing*, whereby sensor patches that provide dense spatial sampling are sparsely placed over a wider area of interest to capture large-scale heterogeneities. These patches however must be interconnected to allow unified data collection and network management. In turn, this need translates to the general *relay placement* problem, described next.

We consider an area  $\mathcal{A}$ , over which  $n$  sensing motes are placed at locations  $\mathcal{S} = \{(x_i^s, y_i^s), i = 1, \dots, n\}$ <sup>4</sup>. Domain scientists determine the sensing locations and therefore we consider them to be fixed [21]. Furthermore, the network has  $m \geq 1$  gateways at locations  $\mathcal{G} = \{(x_j^g, y_j^g), j = 1, \dots, m\}$ . Given  $\mathcal{S}, \mathcal{G}$ , we need to determine the number  $c$  and set of locations  $\mathcal{R} = \{(x_k^r, y_k^r), k = 1, \dots, c\}$  at which to place relay nodes. These nodes use the same radios as the sensing

<sup>4</sup>We do not make any assumptions about the planarity of  $\mathcal{A}$ , but use two dimensions for ease of exposition.

- 
- Step 1** Perform site survey and determine  $\mathcal{S}, \mathcal{G}$ . Measure site-specific channel parameters (§2.1).
  - Step 2** Calculate length of the connected region (§2.2).
  - Step 3** Set effective transmission range  $R_i$  (§3.1).
  - Step 4** Run Steiner tree algorithm, output  $c$  and  $\mathcal{R}$ .
  - Step 5** Deploy relay nodes using Bernoulli Grid (§3.2).
  - Step 6** Post deployment maintenance if necessary.
- 

Table 5. Sensor networks deployment procedure.

nodes to relay data. The relay locations are selected so that every sensing mote has a high quality (i.e., low-loss) end-to-end path to one of the network’s gateways. To do so, we need to ensure that the PRR of every link between the sensing motes and the gateway(s) is above  $p_{high}$ . We note that this model is different from the one proposed by Krause et al. [16] which concurrently optimizes communication and sensing quality by selecting the locations of sensing motes as well as any possible relay nodes. We chose to focus on optimizing the communication topology because this approach matches the domain scientists’ expectation of being able to dictate the sensing locations and because it does not require prior knowledge about the probability distribution of the sensor data.

The relay placement problem is similar to the minimum Steiner Tree problem [14], in which a set of *Steiner points* must be added to create a minimum cost spanning tree of a graph  $G$ . The Steiner Tree problem is NP-Complete even for Euclidean or rectilinear metrics [11]. However, a number of polynomial approximation algorithms have been recently proposed for this problem (see [12, 20] and references therein). These algorithms first compute a Minimum Spanning Tree (MST) covering the nodes in  $\mathcal{S} \cup \mathcal{G}$ . They then determine for each MST link the number and locations of relay nodes necessary based on the link’s length and the relays’ transmission range. These relays are then used to physically connect the two original MST nodes.

These algorithms assume that the transmission range of node  $i$  can be modeled as a disk with known radius  $R_i$ . Unfortunately, selecting an effective transmission radius that ensures low loss is a non-trivial task in practice. One approach is first to estimate the length of the radio’s connected range

Number of Relay Nodes	137	98	91	92	91
Communication Range(m)	30	45	60	75	90

**Table 6. Number of relay nodes needed to connect the network for different communication ranges.**

and then set  $R_i$  to this value. In practice, the length of the connected region depends on the deployment environment and can be derived through an RF site survey. However, this approach will also require redundant relay nodes as the connected range can be as little as 30% to 50% of the radio’s total range [43]. Moreover, placing a relay within the radio’s connected region does not guarantee high PRR (see Figure 2), especially in environments with high RSSI variations.

Setting  $R_i$  within the radio’s transitional area allows the network’s planners to leverage a larger percentage of the radio’s total range, but also raises the risk of adding links with low PRR. Therefore, we can use the proposed Bernoulli Grid method to discover, with high probability, long links with high PRR, at a trivial additional cost. Extensive results presented in Section 4 have shown that the proposed methodology applies to realistic deployment scenarios including the forest shown in Fig.16. Table 5 summarizes the deployment procedures of the network planning tool to apply the Bernoulli Grid method. Note that the Steiner tree algorithm uses each relay to connect at most two nodes, hence the evaluation in Sections 4.6 and 4.7 is sufficient and applicable.

To investigate the improvement achievable for larger networks, we simulate a scenario in which 120 sensing nodes are deployed in a forest spanning  $800m \times 800m$ . The sensing nodes are randomly placed over the deployment area. We then use our planning tool to estimate the number and positions of the required relay nodes. Without the Bernoulli Grid method, one needs to set the maximum communication range to 30 meters, as this is the radio’s connected region in the two forests that we tested. By grid sampling, one can set the effective communication range to  $l_i = 70$  meters at the cost of a few trials when deploying each relay node.

Table 6 summarizes the number of required relay nodes under different effective communication ranges. Two hundred randomly generated network topologies are simulated and the median number of relay nodes is listed in each case. One can see that increasing the effective communication range of relay nodes by leveraging spatial heterogeneity reduces the total number of relay nodes by up to 33.6%. We note that Table 6 serves as an example of the savings achievable for certain types of sensor network deployments. For a small deployment, the places to put relay nodes are straightforward and the formal planning tool shown in Table 5 is not necessary. Nevertheless, the principles of the Bernoulli Grid method are still applicable. We also note that this section assumes that sensor nodes can be placed anywhere, which might not hold due to obstacles and other practical limitations. In these cases, however, one can still try placing a relay node at multiple locations, which do not necessarily form a grid.

The tradeoff inherent to the planning tool is between the quantity of relay nodes and the efforts to deploy them. Al-

lowing more relay nodes would make the deployment process easier as less trials are needed when deploying each node. However, the costs of maintaining those additional relay nodes over the lifetime of the network (up to multiple years) can be much higher than the initial hardware costs of the relay nodes. Therefore, it is up to the network planner to strike the balance. Another cost to use the planning tool is the need for site survey, which is not unique to this tool. In any deployment, some level of site survey is always needed to at least estimate the radio communication range. When using the site planning tool, one can choose to collect measurements to compute the radio propagation parameters, or simply estimate a general communication range and use the Bernoulli Grid method to find good locations.

## 5.2 Mobile Sensor Networks and Sinks

Recent studies have proposed using mobile sinks to harvest data from wireless sensor networks [17, 37]. With the ability to move to different locations, this type of sink can directly benefit from the Bernoulli Grid method: if the current location has low PRR, then the sink can move around the vicinity and try to find a better position. Furthermore, sinks can also utilize the Bernoulli Grid method to purposely construct long links such that the mobile robot that carries the sink can travel a shorter distance, thereby saving the robot’s energy and also accelerating the data collection process.

The Bernoulli Grid method has more interesting implications when more nodes are mobile [38, 40]. Due to the increased mobility, the signal variation in the space domain can be translated to variation in the time domain. For example, when a mobile node sends a packet but does not receive the acknowledgement, it can infer that the current location is probably not a very good one. Therefore, one reasonable action is to wait until the node gets to a different location (e.g., at least 1 meter apart from the previous one) and then try re-sending the packet. The time to wait is determined by the speed of the robot and the separation that is needed to ensure independence. Of course, a mobile robot knows how far it has traveled and therefore does not need to compute the time to wait. However, for a more general class of mobile networks that cannot control their movement, or do not know where they are going, the time to wait can potentially be a useful parameter for communication protocols, assuming some knowledge about the speed of travel is available.

## 6 Related Work

Multiple studies have shown that the characteristics of a network’s wireless links significantly impact upper-level protocols. For example, Ganesan et al. showed that unreliable and asymmetric links cause simple flooding protocols to exhibit unexpected clustering behaviors that differ from the uniform trees generated under ideal radio conditions [10]. Likewise, Woo et al. studied the effects of link connectivity on tree routing protocols for sensor networks [41]. Finally, Zhou et al. found that radio irregularity has a significant impact on geographic routing protocols [44].

This direct connection between wireless link behavior and application-level performance has motivated researchers to measure the properties of low-power wireless links and understand the underlying factors that control their behavior.

Several studies have classified low-power wireless links into three distinct reception regions: *connected*, *transitional*, and *disconnected* [41, 43]. In the connected region, links are of good quality, stable, and symmetric [6]. In contrast, the transitional or *gray* region is characterized by the presence of unreliable and asymmetric links. Unfortunately, the transitional region is often quite large in size and thereby a large number of network links can be unreliable. In fact, Zhao and Govindan found that number to be higher than 50% in the network configurations they tested [43]. Furthermore, the shape of the transitional region is neither circular nor convex [10, 44]. Cerpa et al. also found that there is a significant percentage of the radio range in which links are highly variable, with similar probabilities of having very high or very low reception rates [6].

In a subsequent study, Cerpa et al. measured the temporal properties of low-power wireless links using an indoor testbed of Mica motes [7]. This study found that the quality, in terms of the required number of packets necessary for a successful transmission, of some wireless links can vary considerably over time. However, Cerpa et al. also found that good links tend to be very stable over time [7], an observation we have independently verified.

Son et al. established that signal to interference plus noise ratio (SINR) is the main factor determining PRR in low-power wireless links [33]. Furthermore, the same study found that radio hardware variations affect the SINR threshold necessary for high packet reception, while location does not play a significant role, as long as SINR remains constant [33]. Our results, shown in Figure 1, confirm that when the SNR is greater than a lower bound, the PRR is high with a high likelihood. Otherwise the link enters a gray area where the PRR is difficult to predict. Multiple studies have shown that temporal PRR variations are also due to changes in the signal strength (up to a few dBm) of packets received from the same node over long time periods [18, 34, 36].

Zamalloa and Krishnamachari provided an analysis of the root causes behind link unreliability and asymmetry [42]. Specifically, they used the log-normal path loss model to derive expressions for the distribution, expectation, and variance of the packet reception rate as a function of distance. Rather than using such a model to explain the extent and the location of the transitional region, we leverage it to experimentally find locations within the transitional region that have high PRR, using a small number of trials. We note that the existence of these locations is predicted by analytical models [42] and has been experimentally verified [6, 41, 43]. While we currently do not consider non-isotropic transmission, a characteristic of low-power wireless links [44], it is feasible to incorporate hardware transmission irregularities by introducing variable transmission power  $P_t$  (see Sec.2.1).

Extensive literature covers outdoor [32] and indoor RF signal propagation models, including models that take into account the number, delay, and power of indoor multipath components [1]. We employ the popular log-normal path loss model [28]. This model can be used over large and small distances [31], while empirical studies have shown that it can effectively model multipath indoor channels [22].

Recent efforts provide analytical models for estimating

the connectivity of an ad-hoc network using the log-normal path loss model [3, 13, 19]. These efforts extend earlier work by Bettstetter, who estimated network connectivity under the unit disk model [2]. More recently, Robinson et al. presented a technique for estimating the coverage of metropolitan area WiFi networks using a combination of modeling and analysis [30]. Our work is different because it is *constructive* rather than *descriptive*. Specifically, rather than trying to estimate the connectivity of given networks it provides mechanisms to guide the design of a network by finding locations with high link quality.

## 7 Summary

The Log-normal path loss model has been widely used to describe the propagation of radio signals for 802.15.4 links. In this work, we start with the hypothesis that if the log-normal model holds perfectly, then the signal strengths at nearby locations are independent. In turn, this implies that finding the location whose packet reception ratio is above a certain threshold can be modeled as a sequence of Bernoulli trials. Therefore, the number of attempts required to find a location with high PRR is geometrically distributed. Furthermore, we argue that the geometric distribution is desirable because it does not exhibit long tails and therefore the number of attempts to find a good location should usually be small, provided that a reasonable percentage of location with high PRR exist in the search vicinity.

To verify our hypothesis, we performed extensive experiments in a variety of environments, and concluded that the hypothesis is indeed valid provided that the separation between tried locations is above 1 meter. As a result, we can rely on this hypothesis to perform Bernoulli trials to effectively discover good locations in the transitional region. We showed that this method is useful when deploying sparse sensor networks, as well as for positioning mobile sinks and the nodes of mobile sensor networks.

We also concluded as a rule of thumb that when placing the receiver at a distance that is twice the length of the connected region, one can find a good location in fewer than five trials. This is both practical and efficient, and tells us that finding a good location is usually quite easy. It also serves as the theoretical foundation for the trial-and-error method to deploy sensor motes. While the log-normal model and its application in sensor networks is not new, we argue that this work provides fresh and useful insights into the implications of this model, and is of practical importance.

## 8 Acknowledgments

We extend our gratitude to Neal Patwari, Bhaskar Krishnamachari and Marco Zuniga for their insightful comments and suggestions. We would also like to thank the anonymous reviewers and our shepherd Utz Roedig for helping us improve the paper's presentation. This material is based upon work partially supported by the National Science Foundation under grants #0834470, #0754782, #0546648.

## References

- [1] M. K. Awad, K. T. Wong, and Z. bin Li. An integrated overview of the open literature's empirical data on the indoor radiowave channel's delay properties. *IEEE Transaction on Antennas and Propagation*, May 2008.

- [2] C. Bettstetter. On the connectivity of ad hoc networks. *The Computer Journal*, 2004.
- [3] C. Bettstetter and C. Hartmann. Connectivity of wireless multihop networks in a shadow fading environment. *Wireless Networks*, pages 571–579, 2005.
- [4] C. Boano, N. Tsiftes, T. Voigt, J. Brown, and U. Roedig. The impact of temperature on outdoor industrial sensor network applications. *Industrial Informatics, IEEE Transactions on*, 6:451–459, 2010.
- [5] R. Burns, A. Terzis, and M. Franklin. Software tools for sensor-based science. In *Proceedings of the Third Workshop on Embedded Networked Sensors (EmNets 2006)*, Feb. 2006.
- [6] A. Cerpa, J. L. Wong, L. Kuang, M. Potkonjak, and D. Estrin. Statistical Model of Lossy Links in Wireless Sensor Networks. In *Proceedings of IPSN 2005*, May 2005.
- [7] A. Cerpa, J. L. Wong, M. Potkonjak, and D. Estrin. Temporal Properties of Low Power Wireless Links: Modeling and Implication on Multi-Hop Routing. Technical Report CENS-TR-44, University of California, Los Angeles, Center for Embedded Networked Computing, Jan. 2005.
- [8] Y. Chen and A. Terzis. On the Mechanisms and Effects of Calibrating RSSI Measurements for 802.15.4 Radios. In *Proceedings of the Seventh European Conference on Wireless Sensor Networks (EWSN)*, pages 256–271, Feb. 2010.
- [9] P. Dutta, S. Dawson-Haggerty, Y. Chen, C.-J. M. Liang, and A. Terzis. Design and evaluation of a versatile and efficient receiver-initiated link layer for low-power wireless. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems (Sensys)*, pages 1–14, Nov 2010.
- [10] D. Ganesan, B. Krishnamachari, A. Woo, D. Culler, D. Estrin, and S. Wicker. Complex behavior at scale: An experimental study of low-power wireless sensor networks. Technical Report CSD-TR 02-0013, University of California, Los Angeles, Computer Science Department, Feb. 2003.
- [11] M. Garrey and D. Johnson. The Rectilinear Steiner Problem is NP-Complete. *SIAM Journal of Applied Math*, 13:1351–1365, 1977.
- [12] X. Han, X. Cao, E. L. Lloyd, and C.-C. Shen. Fault-tolerant Relay Node Placement in Heterogeneous Wireless Sensor Networks. In *Proceedings of IEEE Infocom*, 2007.
- [13] R. Hekmat and P. V. Miehem. Connectivity in Wireless Ad-hoc Networks with a Log-normal Radio Model. *Mobile Networks and Applications*, pages 351–360, 2006.
- [14] F. Hwang, D. Richards, and P. Winter. *The Steiner Tree Problem*, volume 53 of *Annals of Discrete Mathematics*. Elsevier, North-Holland, 1992.
- [15] iRobot Corporation. iRobot Create Programmable Robot. Available at <http://www.irobot.com/>.
- [16] A. Krause, C. Guestrin, A. Gupta, and J. Kleinberg. Near-optimal Sensor Placements: Maximizing Information while Minimizing Communication Cost. In *Proceedings of Information Processing in Sensor Networks (IPSN)*, 2006.
- [17] B. Kusy, H. Lee, M. Wicke, N. Milosavljevic, and L. Guibas. Predictive qos routing to mobile sinks in wireless sensor networks. In *Proceedings of the 2009 International Conference on Information Processing in Sensor Networks, IPSN '09*, pages 109–120, Washington, DC, USA, 2009. IEEE Computer Society.
- [18] S. Lin, J. Zhang, G. Zhou, L. Gu, J. A. Stankovic, and T. He. ATPC Adaptive Transmission Power Control for Wireless Sensor Networks. In *Proceedings of the 4th ACM Sensys Conference*, 2006.
- [19] D. Miorandi and E. Altman. Coverage and connectivity of ad hoc networks in presence of channel randomness. In *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, 2005.
- [20] S. Misra, S. D. Hong, G. Xue, and J. Tang. Constrained Relay Node Placement in Wireless Sensor Networks to Meet Connectivity and Survivability Requirements. In *Proceedings of IEEE Infocom*, 2008.
- [21] R. Musáloiu-E., A. Terzis, K. Szlavecz, A. Szalay, J. Cogan, and J. Gray. Life Under your Feet: A Wireless Sensor Network for Soil Ecology. In *Proceedings of the 3rd EmNets Workshop*, May 2006.
- [22] H. Nikookar and H. Hashemi. Statistical modeling of signal amplitude fading of indoor radiopropagation channels. In *2nd International Conference on Universal Personal Communications: Gateway to the 21st Century*, 1993.
- [23] J. D. Parsons. *The Mobile Radio Propagation Channel, 2nd Edition*. John Wiley & Sons, 2000.
- [24] N. Patwari and P. Agrawal. Effects of Correlated Shadowing: Connectivity, Localization, and RF Tomography. In *International Conference on Information Processing in Sensor Networks (IPSN)*, pages 82–93, 2008.
- [25] J. Polastre, J. Hill, and D. Culler. Versatile Low Power Media Access for Wireless Sensor Networks. In *Proceedings of the 2nd ACM Sensys Conference*, 2004.
- [26] J. Polastre, R. Szewczyk, and D. Culler. Telos: Enabling Ultra-Low Power Wireless Research. In *Proceedings of the Fourth International Conference on Information Processing in Sensor Networks: Special track on Platform Tools and Design Methods for Network Embedded Sensors (IPSN/SPOTS)*, Apr. 2005.
- [27] D. Puccinelli and M. Haenggi. Multipath Fading in Wireless Sensor Networks: Measurements and Interpretation. In *IEEE/ACM International Wireless Communications and Mobile Computing Conference (IWCMC'06)*, Vancouver, Canada, July 2006. Available at <http://www.nd.edu/~mhaenggi/pubs/iwcmc06.pdf>.
- [28] T. S. Rappaport. *Wireless Communications: Principles & Practices*. Prentice Hall, 1996.
- [29] N. Reijers, G. Halkes, and K. Langendoen. Link layer measurements in sensor networks. In *1st IEEE Int. Conf. on Mobile Ad-hoc and Sensor Systems*, oct 2004.
- [30] J. Robinson, R. Swaminathan, and E. Knightly. Assessment of Urban-Scale Wireless Networks with a Small Number of Measurements. In *Proceedings of Mobicom 2008*, Sept. 2008.
- [31] S. Seidel and T. Rappaport. 914 MHz path loss prediction model for indoor wireless communication in multi floored buildings. *IEEE Transactions on Antennas and Propagation*, pages 207–217, 1992.
- [32] K. Sohrawi, B. Manriquez, and G. Pottie. Near ground wideband channel measurement in 800-1000 MHz. In *Proceedings of the IEEE 49th Vehicular Technology Conference*, 1999.
- [33] D. Son, B. Krishnamachari, and J. Heidemann. Experimental study of concurrent transmission in wireless sensor networks. In *Proceedings of the ACM Sensys*, Nov. 2006.
- [34] K. Srinivasan, P. Dutta, A. Tavakoli, and P. Levis. Some implications of low power wireless to IP networking. In *Proceedings of the Fifth Workshop on Hot Topics in Networks (HotNets-V)*, Nov. 2006.
- [35] K. Srinivasan, M. Jain, J. I. Choi, T. Azim, E. S. Kim, P. Levis, and B. Krishnamachari. The  $\kappa$ -factor: inferring protocol performance using inter-link reception correlation. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking, MobiCom '10*, pages 317–328, New York, NY, USA, 2010. ACM.
- [36] K. Srinivasan, M. Kazandijeva, S. Agarwal, and P. Levis. The  $\beta$ -factor: Measuring Wireless Link Burstiness. In *Proceedings of the 6th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, 2008.
- [37] O. Tekdas, J. H. Lim, A. Terzis, and V. Isler. Using Mobile Robots to Harvest Data from Sensor Fields. *IEEE Wireless Communications Magazine*, 26:22–28, Feb. 2009.
- [38] O. Tekdas, W. Yang, and V. Isler. Robotic routers: Algorithms and implementation. *Int. Journal of Robotics Research*, 29(1), 2010.
- [39] Texas Instruments. 2.4 GHz IEEE 802.15.4 / ZigBee-ready RF Transceiver. Available at <http://www.ti.com/lit/gpn/cc2420>, 2006.
- [40] P. Tokekar, D. Bhaduria, A. Studenski, and V. Isler. A robotic system for monitoring carp in minnesota lakes. *Journal of Field Robotics*, 2010. accepted.
- [41] A. Woo, T. Tong, and D. Culler. Taming the underlying challenges in reliable multihop wireless sensor networks. In *Proceedings of 1st ACM Sensys Conference*, 2003.
- [42] M. Z. Zamalloa and B. Krishnamachari. An analysis of unreliability and asymmetry in low-power wireless links. *ACM Transactions on Sensor Networks*, June 2007.
- [43] J. Zhao and R. Govindan. Understanding Packet Delivery Performance in Dense Wireless Sensor Networks. In *Proceedings of the ACM Sensys*, Nov. 2003.
- [44] G. Zhou, T. He, S. Krisnamurthy, and J. Stankovic. Impact of radio irregularity on wireless sensor networks. In *Proceedings of ACM Mobisys 2004*, 2004.