Report Number: wucse-2009-26

2009

Radio Mapping for Indoor Environments

Octav Chipara, Gregory Hackmann, Chenyang Lu, and William D. Smart

The efficient deployment and robust operation of many sensor network applications depend on deploying relays to ensure wireless coverage. Radio mapping aims to predict network coverage based on a small number of link measurements from sampled locations. Radio mapping is particularly challenging in complex indoor environments where walls significantly affect radio signal propagation. This paper makes the following key contributions to indoor radio mapping. First, our empirical study in an office building identifies a wall-classification model as the most effective model for indoor environments due to its balance between model complexity and accuracy. Second, we propose a practical algorithm... **Read complete abstract on page 2.**

Radio Mapping for Indoor Environments

Octav Chipara, Gregory Hackmann, Chenyang Lu, and William D. Smart

Complete Abstract:

The efficient deployment and robust operation of many sensor network applications depend on deploying relays to ensure wireless coverage. Radio mapping aims to predict network coverage based on a small number of link measurements from sampled locations. Radio mapping is particularly challenging in complex indoor environments where walls significantly affect radio signal propagation. This paper makes the following key contributions to indoor radio mapping. First, our empirical study in an office building identifies a wall-classification model as the most effective model for indoor environments due to its balance between model complexity and accuracy. Second, we propose a practical algorithm to predict the Reception Signal Strength (RSS) of links in an indoor environment based on a small number of measurements at sampled locations. A key novelty of our algorithm lies in its capability to automatically classify walls into a small number of classes with different degrees of signal attenuation. Finally, we present a practical Radio Mapping Tool that can predict the coverage areas of relays based on a small number of link quality measurements in the environment. Empirical evaluation in an office building demonstrates that the Radio Mapping Tool reduces the false positive rate by as much as 41% compared to the classical log-normal radio propagation model, with a false negative rate of 9% based on sampling only 20% of the locations of interest.

2009-26

Radio Mapping for Indoor Environments

Authors: Octav Chipara, Gregory Hackmann, Chenyang Lu, William D. Smart, Gruia-Catalin Roman

Abstract: The efficient deployment and robust operation of many sensor network applications depend on deploying relays to ensure wireless coverage. Radio mapping aims to predict network coverage based on a small number of link measurements from sampled locations. Radio mapping is particularly challenging in complex indoor environments where walls significantly affect radio signal propagation. This paper makes the following key contributions to indoor radio mapping. First, our empirical study in an office building identifies a wall-classification model as the most effective model for indoor environments due to its balance between model complexity and accuracy. Second, we propose a practical algorithm to predict the Reception Signal Strength (RSS) of links in an indoor environment based on a small number of measurements at sampled locations. A key novelty of our algorithm lies in its capability to automatically classify walls into a small number of classes with different degrees of signal attenuation. Finally, we present a practical Radio Mapping Tool that can predict the coverage areas of relays based on a small number of link quality measurements in the environment. Empirical evaluation in an office building demonstrates that the Radio Mapping Tool reduces the false positive rate by as much as 41% compared to the classical log-normal radio propagation model, with a false negative rate of 9% based on sampling only 20% of the locations of interest.

Type of Report: Other

Radio Mapping for Indoor Environments

Octav Chipara, Gregory Hackmann, Chenyang Lu, William D. Smart, Gruia-Catalin Roman
Department of Computer Science and Engineering
Washington University in St. Louis, USA
{ochipara, gwh2, lu, wds, roman}@cse.wustl.edu

Abstract

The efficient deployment and robust operation of many sensor network applications depend on deploying relays to ensure wireless coverage. *Radio mapping* aims to predict network coverage based on a small number of link measurements from sampled locations. Radio mapping is particularly challenging in complex indoor environments where walls significantly affect radio signal propagation. This paper makes the following key contributions to indoor radio mapping. First, our empirical study in an office building identifies a wall-classification model as the most effective model for indoor environments due to its balance between model complexity and accuracy. Second, we propose a practical algorithm to predict the Reception Signal Strength (RSS) of links in an indoor environment based on a small number of measurements at sampled locations. A key novelty of our algorithm lies in its capability to automatically classify walls into a small number of classes with different degrees of signal attenuation. Finally, we present a practical Radio Mapping Tool that can predict the coverage areas of relays based on a small number of link quality measurements in the environment. Empirical evaluation in an office building demonstrates that the Radio Mapping Tool reduces the false positive rate by as much as 41% compared to the classical log-normal radio propagation model, with a false negative rate of 9% based on sampling only 20% of the locations of interest.

1 Introduction

Numerous sensor network applications require wireless networks which fully cover an entire physical region. Examples of such applications include participatory sensing [4,8], elderly care [25,28], and patient tracking and monitoring [3]. Our own interest in this topic is motivated by a medical application which requires covering a hospital floor to ensure that the vital signs of mobile users may be collected reliably. Accordingly, relays should be deployed such that, for any position where a patient may move, there is at least one link to some relay that exceeds a minimum bound on packet reception rate (PRR). The best practice for assessing the coverage of such a sensor network is to exhaustive measure the link quality to the deployed relays at numerous locations. This process is labor intensive and leads to significant deployment costs. Even worse, the network coverage may need to be reassessed in response to physical changes (e.g., reconfiguring cubicles) or changes in the radio properties (e.g., switching operating frequency due to interference). As a result, maintenance cost may also be significant.

Within the 802.11 networking community, there are a handful of deployment tools which can proactively evaluate radio coverage [5]. These tools often employ ray tracing techniques which require precise characterizations of the location and radio properties of walls. While there are empirical studies which determine the attenuation of different wall types [21], the user must manually determine the construction material of each wall. Such information is seldom readily available to the network manager or end user. The user may also need to provide the locations and attenuation coefficients of other objects that can significantly affect radio propagation, such as bookshelves or filing cabinets. This imposes an undue effort on the user.

What is needed is a *practical radio mapping tool* which reduces the burden of assessing the coverage of a wireless sensor network based on a small number of link measurements. Specifically, we are interested in determining the reception coverage of a relay: i.e., the set of points (x,y) such that nodes located at (x,y) can transmit a packet to at least one relay with a PRR above a user-specified threshold[1]. The total network coverage may be computed by taking the union of each relay's coverage. We choose to focus on ensuring a minimum PRR because PRR directly affects the performance of link-layer and routing protocols.

We divide the radio mapping problem into two parts. We first predict the receive signal strength (RSS) at the relay

SenSys'09, November 4–6, 2009, Berkeley, CA, USA.

[1] The techniques proposed in this paper are also applicable to the network's transmission coverage: i.e., the set of points that can receive transmissions from at least one relay. We focus on reception coverage in this paper, since our target application entails data collection. Henceforth, we use the term "coverage" to mean "reception coverage".

from any point on the floor plan. Then, based on the RSS predictions along with an RSS threshold for predicting good-quality links, we determine each relay's coverage. While established radio propagation models have been proposed to predict RSS [1, 10, 12], it is unclear which is best-suited for complex indoor environments where nodes may not have line-of-sight. Finding an RSS threshold that predicts good-quality links is also challenging because low-power links have probabilistic properties [7, 15, 17, 22, 23].

The key contributions of this paper are the following:

- We perform an in-depth *empirical study* that characterizes the accuracy of several propagation models' RSS predictions in an office building. In this study, we show that complex models do not necessarily produce more accurate estimates of signal strength: there is an important tradeoff between the accuracy of the model and the number of model parameters that must be estimated from limited training data. We demonstrate that a model which classifies walls into a small number of classes with differing attenuation values achieves the best prediction accuracy, reducing errors by 18% compared to the classical log-normal radio propagation model [2].
- We propose a computationally efficient *wall classification algorithm* to be used in conjunction with this wall-classification-based model. Our empirical results show that it produces a wall classification which results in better RSS prediction accuracy than a manual classification, while also significantly reducing the user burden.
- We develop a practical *Radio Mapping Tool* (RMT) which predicts the coverage of one or more relays with minimal information about the indoor environment. RMT leverages wall location information to accurately predict radio coverage based on a small number of link quality measurements. Moreover, it can compute the predicted coverage of a network in minutes, allowing users to quickly refine the placement of relay nodes.
- We characterize the accuracy of this tool through a *case study*. We find that the combination of our chosen radio model with our automated wall classification scheme reduces the false positive rate by as much as 41% compared to the log-normal model, with a false negative rate of 9% based on sampling only 20% of the locations of interest.

The remainder of the paper is organized as follows. In Section 2, we start by overviewing several well-established propagation models and discussing their applicability to indoor environments. We find a model which classifies walls into a few types to be promising. In Section 3, we discuss methods to classify walls, including a computationally efficient algorithm that automatically performs this classification. The RSS prediction accuracy of different propagation models is assessed in Section 4. In Section 5, we present a radio mapping tool built based on the insights gained from our empirical study. Section 6 evaluates the efficacy of our radio mapping tool through a case study. We review related work Section 7 and conclude in Section 8.

2 Radio Propagation Models

Modeling signal propagation in wireless networks has attracted tremendous interest within the wireless communication community. Models optimized for different wireless technologies and environments have been proposed in literature [1, 10, 12]. The models presented in this section focus on modeling three aspects which may significantly affect signal propagation in indoor environments: (1) the distance between the sender and receiver, (2) variations in the transmission power caused by imperfect radio calibration, and (3) the impact of walls. While these models are not new, they have not been systematically evaluated on low-power radios in complex indoor environments where nodes may not have line-of-sight. Our goal is to identify the model with the best trade-off between prediction accuracy and the number of samples needed to estimate its parameters. This model will ultimately be used in our Radio Mapping Tool to generate signal strength predictions. As we show in Section 4, more complex models do not necessarily result in improved prediction accuracy: it may not be possible to accurately estimate many parameters with a reasonable number of samples.

Log-Normal Shadowing: Under the log-normal model, signal strength decays exponentially as a function of distance. Let $d(s,r)$ be the distance between the sender node s and the receiver node r. The receive signal strength $P_r(s,r)$ at r from a sender s is given by: [2]

$$P_r(s,r) = \alpha - 10\beta \log_{10} d(s,r) + \sigma \quad (1)$$

The parameter α represents the transmission power at a reference distance of 1m. β represents the pass loss exponent. σ models shadowing (i.e., the random signal variations between sender and receiver) and is usually considered to be a normally distributed random variable.

An attractive feature of the log-normal model is its simplicity: it has only two parameters that may be determined through linear regression on a small number of samples. Moreover, we may account for the variation in transmission power among radios caused by imperfect radio calibration [7, 22, 30] by having different α parameters for each sender (i.e., replacing α in Equation 1 with $\alpha(s)$). This model has been shown in [30] to explain the wide transitional region observed on low-power links.

Prior empirical studies have shown that this model may accurately predict the receive signal strength of low-power radios in outdoor environments [15] and in indoor environments where nodes have direct line-of-sight [15, 30]. Since this model does not account for the impact of walls, we do not expect it to perform well in complex indoor environments; the results of our empirical study in Section 4 will later confirm this hypothesis.

Sector-Based: Prior literature has extended the basic log-normal model to capture the fact that many low-power radios, such as the CC2420 radio used in the TelosB and MicaZ sensor platforms, have non-isotropic radiation patterns [29]. That is, even when nodes are positioned at equal distances from the sender, they may observe significantly different receive signal strengths. Intuitively, by capturing this effect we obtain a more realistic model which should result in more ac-

curate signal strength predictions.

This extension captures the effect of non-isotropic radiation by parameterizing α by the angle θ between the line uniting s and r and a fixed frame of reference:

$$P_r(s,r) = \alpha(s,\theta) - 10\beta\log_{10} d(s,r) + \sigma \quad (2)$$

$\alpha(\theta)$ may be a non-linear function [29]. As a result, non-linear optimization techniques would be necessary for fitting the model. A simpler alternative is to divide θ into NS sectors:

$$P_r(r,s) = \sum_{j=1}^{NS} B_{(1+\lfloor\theta/NS\rfloor=j)}\alpha(s,j) - 10\beta\log_{10} d(s,r) + \sigma \quad (3)$$

where $B_{(a=b)}$ is one if $a = b$ and zero otherwise. The parameter values of each sector may then be fitted independently using linear regression.

While this model is more realistic than the basic log-normal model, it is also more complex. Accordingly, it will require much more data to adequately fit its parameters. Under the log-normal model, each sender has a single value for α; under the sector-based model, each sender has NS values of α. In our testing environment (shown in Figure 2), this model may have as many as 360 parameters. In Section 4 we show that the large number of parameters causes the sector-based model to perform poorly in our indoor environment.

Per-Wall Attenuation: In complex indoor environments, walls may significantly attenuate wireless links. We hypothesize here (and validate in Section 4) that incorporating walls into the radio propagation model can improve its signal strength predictions. Such information can be readily extracted from a building's floor plan.

An intuitive way of modeling wall attenuation is to assume that each wall $w \in W$ in the environment attenuates the signal by a constant factor γ_w. Therefore, if we let $I_{s,r}$ be the set of all walls which intersect a virtual line between s and r, then the signal strength at r is given by:

$$P_r(s,r) = \alpha(s) - 10\beta\log_{10} d(s,r) + \sum_{w\in I_{s,r}} \gamma_w \quad (4)$$

This model may also be combined with the non-isotropic radio range model, in which case α is replaced with $\sum_{j=1}^{NS} B_{(1+\lfloor\theta/NS\rfloor=j)}\alpha(s,j)$.

However, several measurements should be taken through each wall to accurately estimate γ. This may be a significant burden in some environments; for example, we identified 85 walls in our testing environment. We demonstrate in Section 4 that this requirement causes the per-wall attenuation model to make large estimation errors when only a small amount of training data is available. In some cases, this model performs even worse than the log-normal model.

Wall-Class Attenuation: A pragmatic alternative to the per-wall scheme is to group walls into a few classes, reflecting the fact that only a few types of walls are used in construction. For example, the building shown in Figure 2 mainly uses two kinds of walls: cinder block and drywall. Given a set of classes C, a mapping $\Pi : W \to C$, and an attenuation coefficient Γ_c for each class $c \in C$, the signal strength at a node r is:

$$P_r(s,r) = \alpha(s) - 10\beta\log_{10} d(s,r) + \sum_{w\in I_{s,r}} \Gamma_{\Pi(w)} \quad (5)$$

A benefit of considering a small number of classes is that it significantly reduces the number of parameters. However, it also creates a new problem: a mapping Π from walls to classes needs to be constructed. In the next section, we will discuss an efficient algorithm for automatically constructing this mapping.

3 Estimating Model Parameters

For most of the models discussed in the previous section, the parameter values may be estimated through linear regression as follows. The user collects a training set of link quality measurements by placing nodes at a small number of locations in the environment. The nodes exchange packets and record the receive signal strength (RSS) and sequence number of each packet that they successfully decode. The user also records the location of each deployed node so that the distance d between any two nodes may be computed. From this training data, we compute a vector y which specifies the average signal strength at each node. We can then use standard linear regression techniques to fit most of the models' parameters from d and y.

However, this approach is not sufficient for models based on wall classification, which also require a mapping Π from walls to wall classes. We note that these models are particularly important for our study, because (as discussed in Section 4) they achieve the most accurate predictions of signal strength. Thus, in this section we will focus on an efficient method for fitting parameters to this model.

One way to construct this wall classification is to manually classify walls based on their construction material. Linear regression may then be used to fit the remainder of the model's parameters as described above. However, this approach is sub-optimal for two reasons. First, manual wall classification is labor-intensive and requires architectural information that may not be readily available to application developers or network managers. Second, a manual mapping will not consider the attenuation of other objects in the building. For example, drywall is known to have a lower attenuation than cinderblock. However, a signal that intersects a bookshelf placed against drywall may have an attenuation closer to cinderblock, which a manual architectural survey will not capture.

Instead, we propose to automatically classify walls using an iterative algorithm. This algorithm takes as input the average RSS vector y, a set of walls W, a set of wall classes C, and the distance among all nodes d. Figure 1 presents the pseudocode of this algorithm.

Initially, each wall is assigned to a random class. The algorithm then proceeds in two stages, repeating until changes in wall classification stop improving the sum of squared errors (SSE) between the predicted signal strengths ($\hat{y}$) and the actual signal strengths (y). In the first stage (line 6), the algorithm uses linear regression to fit the parameters α and β,

$[\alpha, \beta, \Gamma, \Pi]$ = compute-parameters(y, W, C, d):

1: improvement = true;
2: for each wall $w \in W$:
3: $\Pi(w) = rand(C)$;
4: while (improvement):
5: improvement = false;
6: $[\alpha, \beta, \Gamma] = regress(y, [d; \Pi])$;
7 : for each wall $w \in W$ in random order:
8: $\Pi_{new} = \Pi$ and $c_{old} = \Pi(w)$;
9: for each class $c \in C$:
10: $\Pi_{new}(w) = c$;
11: $\hat{y} = \alpha(s) - 10\beta \log_{10} d(s,r) + \sum_{w \in I_{s,r}} \Gamma_{\Pi_{new}(w)}$;
12: $SSE(c) = \sum_{i=1}^{|y|} (y(i) - \hat{y}(i))$;
13: $c_{best} = arg\,min_c SSE(c)$;
14: if $(c_{old} \neq c_{best})$:
15: $\Pi(w) = c_{best}$;
16: improvement = true;
17: break;

Figure 1. Fitting algorithm for wall-classification models

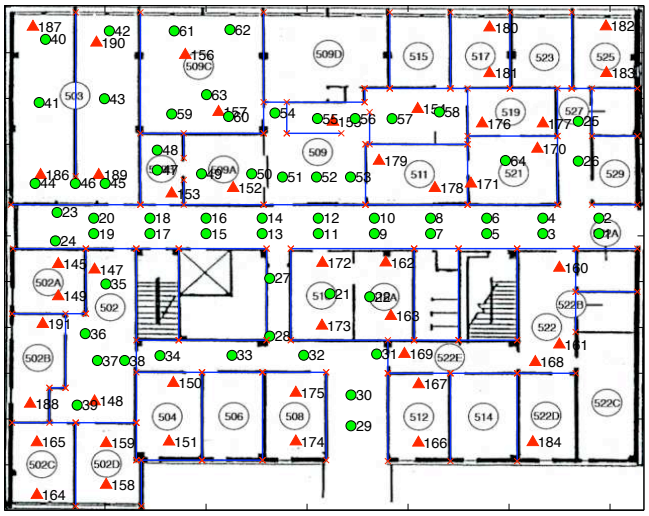


Figure 2. Testbed configuration; triangles represent relays and circles represent test positions

as well as the attenuation coefficient Γ for each wall class. In the second stage (lines 7–16), the algorithm attempts to improve the mapping of walls to classes while keeping the values of α, β, and Γ fixed. This is done by considering each wall w in random order and computing the SSE when w is assigned to each class in C. If changing w's assignment results in a smaller SSE, then w's classification is updated accordingly and the algorithm goes back to executing the first stage with an improved classification of walls. Otherwise, w is already assigned to the best class, and the algorithm considers the next wall. The algorithm terminates when no wall may be assigned to a new class that reduces the SSE. At the completion of the algorithm, the values of the parameters α, β, and Γ are returned along with the mapping Π of walls to classes.

This algorithm has two noteworthy features. First, it is much less computationally expensive that an exhaustive search. The wall-reassignment stage considers at most $|C| \times |W|$ potential assignments at each iteration. We found that, by randomizing the order in which walls are considered, improved assignments are typically found after considering only a few walls. Thus, in practice, this algorithm could be executed in under two minutes on a modern laptop PC.

Second, the algorithm is guaranteed to converge. This is because the algorithm reduces the squared error at each step until it terminates. There is no guarantee on the optimality of the solution, since it may get stuck in a local minimum; however, we show in Section 4 that the classifications found using this algorithm result in better prediction accuracy than a manual wall classification. In fact, because of the random initial assignment of walls to classes and the random ordering in which walls are reclassified, the algorithm may return different values each time it is run. Accordingly, we may improve the squared error by repeating the algorithm several times and returning the parameters which resulted in the lowest squared error.

4 Empirical Study of Signal Strength Models

In this section, we present an empirical study which evaluates the signal propagation models discussed in Section 2. The results of our study indicate that the automatic wall classification model achieves the most accurate signal strength predictions out of all the methods discussed in Section 2. Perhaps more importantly, the findings highlight the importance of selecting a model with the right amount of complexity. The manual and automatic wall classification models achieve as much as 16% lower error than the log-normal model, because the log-normal model is excessively simplistic. However, the sectorization and per-wall attenuation models produce significantly *higher* error than the simpler wall classification models, indicating that a more complex model does not always produce more accurate results. These observations serve as the foundation for developing our radio mapping tool, which we will discuss in Section 5.

4.1 Experimental Setup

Our experiments are carried out on indoor office building (shown in Figure 2) using TelosB motes. The motes are equipped with CC2420 low-power radio chips, which provide an RSS indicator reading for each correctly decoded packet with an accuracy of ± 6 dB [27]. All nodes in our experiment were set to 802.15.4 channel 26, which does not overlap with the 802.11g network deployed in the same building.

The experimental setup is motivated by our interest in supporting robust data collection from mobile users. Accordingly, we are interested in ensuring that at least one testbed node is capable of receiving data from a user standing in any location. A testbed of 45 TelosB motes deployed close to the ceiling is used to represent potential relay locations. As shown in Figure 2, our testbed does not have direct line-of-sight among many nodes. Our study consists of 2880 link quality measurements taken from 64 different positions (circles in Figure 2) to each of the 45 relays (triangles in Figure 2). At each position, a TelosB transmitter broadcasts 800

packets. The relays record the RSS indicator reading and sequence number of each packet which they successfully decode. This information is relayed to a central database using a USB and Ethernet backbone, so as not to interfere with wireless transmissions.

After collecting the data, we randomly selected a portion of the nodes from each room or hallway to use as training data. This training data was input into several different radio propagation models. We then used the models discussed in Section 2 to predict the average signal strength observed over the links which were not represented in the training data, using the remainder of the experimental dataset as ground truth data for comparison. We varied the proportion of training data from 20%–80% during our evaluation.

Unless mentioned otherwise, the presented results are averages of 10 randomly generate training sets. The error bars in all graphs represent the 90% confidence intervals across the 10 runs. (We note that the error bars are sometimes smaller than the markers in the graphs, and hence are not always visible.)

4.2 Effect of Walls

First, we will evaluate the effectiveness of including walls in our radio propagation models by comparing the performance of the two wall-based models against the log-normal model. We also evaluate the efficacy of the automatic wall classification algorithm discussed in Section 3; to do so, we include an additional experimental run using a wall classification that has been manually generated from architectural information of the building.

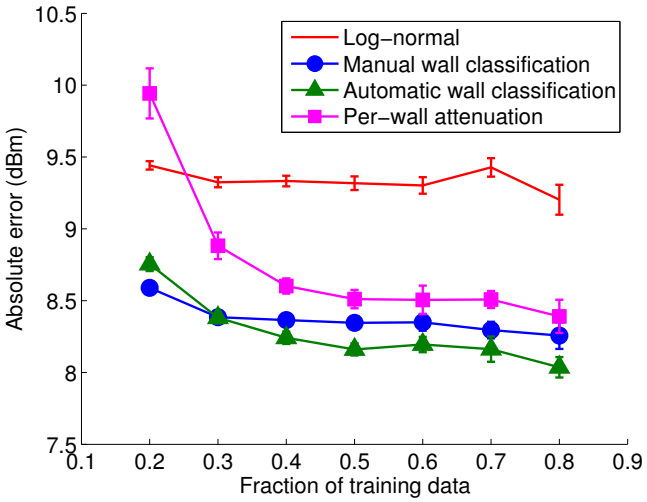


Figure 3. Comparison of radio propagation models

Figure 3 compares the 90th percentile errors of all four experiments with various proportions of training data. The two wall-classification models consistently outperform the log-normal model, which has a prediction error 7.9%–15.5% higher than the automatic wall-classification model and 9.9%–13.7% higher than the manual wall-classification model. The log-normal model is largely unimproved even when it is given substantially more training data; quadrupling the size of the training dataset only decreases the estimation error by 2.5%.

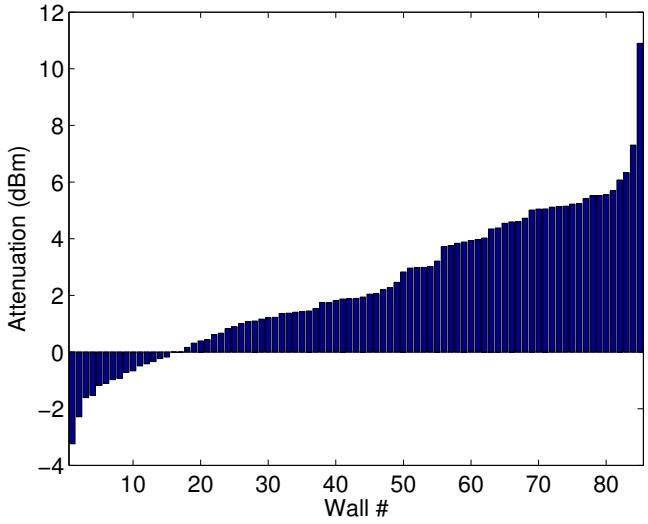


Figure 4. The estimated attenuation of each wall under the per-wall attenuation model

Strikingly, the model which assigns an individual attenuation to each wall has 3.8%–13.6% higher error than the wall-class model with manual classification, and 1.6%–15.8% higher error than the wall-class model with automatic classification. When given a small proportion (20%) of training data, it even performs 5.3% worse than the basic log-normal model that completely ignores wall attenuation, though it improves on the basic model by 5.0%–10.8% when given more training data. As hypothesized in Section 2, this occurs because the per-wall attenuation model has 87 parameters, which cannot be fit well given a reasonable amount of training data. This effect can be seen in Figure 4, which plots the estimated attenuation of each wall in the building; even when 80% of the dataset is used as training data, the model predicts that 17 of the 85 walls *amplify* radio signals.

We also note that the automatic wall-classification scheme achieves comparable error to the manual wall-classification scheme. When both models are given only 20% of the data for training, the automatic wall-classification scheme has 1.9% higher error than the manual wall-classification scheme; at all other datapoints, the automatic wall-classification scheme achieves 0.1%–2.7% lower error than the manual scheme. This occurs because the manual classification scheme only captures the effect of walls, while the automatic scheme can indirectly capture the effect of other obstacles in the environment. However, when the proportion of training data is 20%, the automatic classification scheme does not have enough data to adequately map walls to classes.

From this data, we make the following key insights:

- The basic log-normal radio model can be significantly affected by obstacles; it is not always practical to compensate for this effect by just collecting more training data.
- Judicious use of wall location information can lead to significant improvements in prediction accuracy.

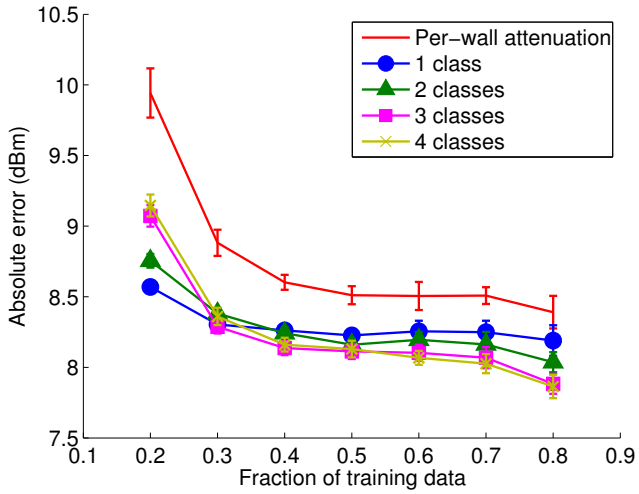


Figure 5. Comparison of different numbers of wall classes

- Automatic wall classification obviates the need to manually classify walls, with comparable or slightly improved prediction accuracy.

4.3 Wall Classification Methods

The previous experiment assumed that all walls in the environment can be assigned into 2 classes. We will now take a closer look at how altering the number of wall classes can affect the accuracy of the automatic wall-class attenuation model.

Figure 5 compares the estimation accuracy of the automatic wall-classification model with 1–4 classes of walls. For comparison, we also include the per-wall model, where each wall is effectively assigned to its own class. When relatively little training data is available, using fewer wall classes improves the predication accuracy; when 20% of the data is used for training, the 1-class model has 2.2%–6.7% lower error than the other classification-based models, and 16.0% lower error than the per-wall model. As the amount of training data is increased, schemes which use more wall classes are able to better fit their data to the model's parameters, and their estimations improve. When the proportion of training data is 30%, the difference among the wall-classification schemes is insignificant; as the proportion is increased to 80%, the 4-class model achieves up to 4.1% lower error than the other class-based models. Nevertheless, at all data points, using 1–4 wall classes consistently improves (by a margin of up to 16.0%) on the most extreme case where each of the 85 walls has its own class.

We thus make the following new key insight into the behavior of the automatic wall-classification model:

- Classifying walls into a small number of classes achieves the lowest error when little training data is available; but as the amount of training data increases, more classes should be employed.

4.4 Impact of Sectorization

In the preceding models, we have ignored the effect of the non-uniform radiation that has been observed on the CC2420 radio chip. We will now explore the sectorization technique (Equation 3) that aims to improve the accuracy of the radio propagation model by modeling this non-isotropic radiation pattern.

Figure 6(a) compares the quality of the signal strength estimates of the automatic wall-classification based model with enhanced models which divide the radio range into 4, 6, and 8 sectors and assign a different reference transmission power (α) to each range. (In this figure, walls are classified into 3 classes; the results for different numbers of classes are similar.) Notably, when 20–70% of the data is used for training, the three schemes which perform sectorization generate 0.3%–71.0% higher errors than a model which does not use sectors; the errors increase as more sectors are added. When 80% of the data is used for training, the differences among the schemes are statistically insignificant. Sectorizing the path-loss exponent (β) instead gives similar results (see Figure 6(b)).

This phenomenon is caused by the fact that modeling non-isotropic communication ranges greatly increases the number of parameters to the model. Rather than solving for a single parameter α or β, it now necessary to solve for up to 360 values of α or β (45 nodes $\times$ 8 sectors/node); much more training data is needed to adequately fit these added parameters. Moreover, the effects of wall attenuation dominate the effects of non-isotropic radiation in our complex indoor environment. Thus, while adding sectorization makes the resulting model more realistic in principle, in practice it does not result in better RSS predictions.

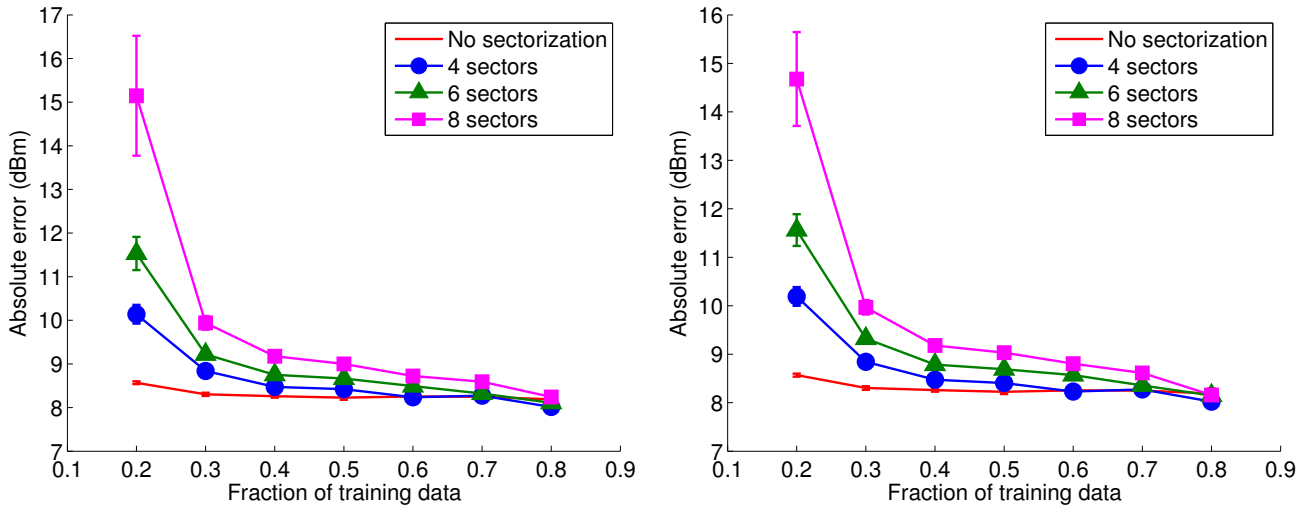
We therefore make one more key insight:

- More sophisticated models are not necessarily better; the difficulties of fitting the extra parameters may outweigh the enhanced model accuracy.

5 A Radio Mapping Tool

In this section, we present our Radio Mapping Tool (RMT), which is designed to assess network coverage. RMT is particularly beneficial for applications which require a network to cover an entire physical area. Examples of such applications include elderly care [25, 28] and the collection of vital signs from ambulatory patients in hospitals [13]. The main uses case of RMT is to determine the coverage of an already deployed network. RMT does this by predicting the coverage of each relay node. The network coverage is then computed as the union of the regions covered by each relay. Coverage gaps are detected by checking if there is any point which is not covered by any relay. RMT may also be used for network deployment. In this case, the user may deploy an overly dense network of relays. Based on the coverage predictions made by RMT, the smallest subset of relays necessary to cover an area may be determined.

RMT has several salient features. (1) In contrast to ray-tracing techniques, RMT does not require the user to specify the attenuation coefficients or construction materials of walls. Wall locations may be extracted from readily available floor plans. (2) At its core, RMT incorporates the insights found in the previous section. Accordingly, RMT uses the models which involve a small number of wall types since they have been found to provide the best trade-off be-



(a) Sectorizations of α

(b) Sectorizations of β

Figure 6. Comparison of sectorization schemes

tween model complexity, number of samples required for accurate parameter estimation, and prediction accuracy. (3) RMT uses the computationally efficient algorithm presented in Section 3 to classify each wall in a small number of classes and determine the model parameters.

RMT requires the user to provide: the locations of the relay (or group of relays) whose coverage the user is interested in predicting, the locations of wall, a training set, and a PRR threshold (PPR_t) used to discriminate between "good" and "bad" links. The training set includes a number of link quality measurements obtained by broadcasting packets at a small number of sample locations. For each link in the training set, the following information must be supplied: the locations of the sender and receiver, the received the packets, and a table of raw link statistics. The raw link statistics include the RSS for each received packet and its sequence number. The sequence numbers are used to compute the PRR of each link. A TinyOS application is include with RMT to assist with this process. RMT then computes summary statistics for each link which include the their PRR and average RSS.

RMT has three main components: a *Parameter Estimator*, an *RSS Mapper*, an *RSS-to-PRR Mapper*, and a *Coverage Mapper* (see Figure 8). The *Parameter Estimator* uses the computationally efficient algorithm described in Section 3 is used to determine the parameters of this model. Based on the determined parameter values, the *RSS Mapper* constructs RSS map which includes signal strength predictions on a dense 2D grid overlaid on the floor plan. The *RSS Mapper* uses the automatic wall classification model. For each relay whose coverage the user wants to predict, the *RSS Mapper* constructs an RSS map. The radio map contains the predicted RSS at the considered relay from a sender that may be placed at each grid point.

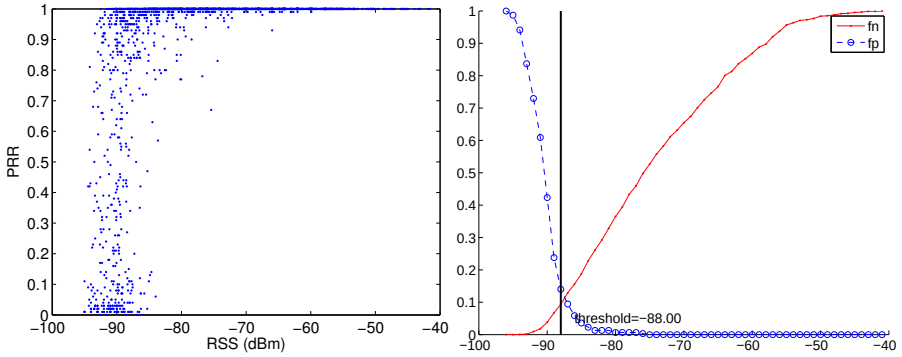
The *RSS-to-PRR Mapper* is responsible for identifying an RSS threshold which accurately separates the "good links" (with PRR higher than PRR_t) from the "bad links" (with PRR lower than PRR_t). An ideal RSS threshold should result in as few false negatives (links incorrectly predicted as good-quality) and false positive (links incorrectly predicted as poor-quality) as possible. Figure 7 illustrates the relationship between RSS and PRR observed in our testing environment. Each point in the figure plots the average RSS of a link against that its PRR. This scatter plot shows that there is a transitional region at approximately -95 to -85 dBm. Our results agree with the in-depth analysis of the correlation between RSS and PRR presented in [24].

To understand how to determine an appropriate RSS threshold let us consider the case when the user specifies good links to have a PRR threshold higher than 80% . For this case, we plot the false negative and false positive rates for each possible RSS threshold between the minimum and maximum observed RSS values. These results are plotted in Figure 7(b). As expected, increasing the RSS threshold decreases the false positive rate while increasing the false negative rate. Setting the RSS threshold to -88 dBm offers the good tradeoff between these rates: the false negative rate is 9% and false positive rate is 10%.

RSS-to-PRR Mapper automatically identifies the appropriate RSS threshold for a user-specified PRR threshold. This is done by having the user provide bounds on the maximum tolerable false negative and false positive rates. RMT selects the minimum RSS which satisfies both bounds, or reports an error to the user when no such RSS threshold exists.

Finally, the *Coverage Mapper* determines the coverage prediction at is grid location by comparing the predicted RSS with the RSS threshold. Figures 9 and 10 are examples of the output maps produced by RMT. We found this to be an effective way of visualizing RMT's signal strength and coverage predictions.

The initial RMT prototype had some performance limitation. We found that the key to ensuring RMT's performance was to precompute as much of the data as possible. For example, we precompute the set of walls which are intersected by the line between any relay and any grid location. This



(a) RSS vs. PRR scatter plot (b) False negative and negative rates

Figure 7. Selecting an appropriate RSS threshold

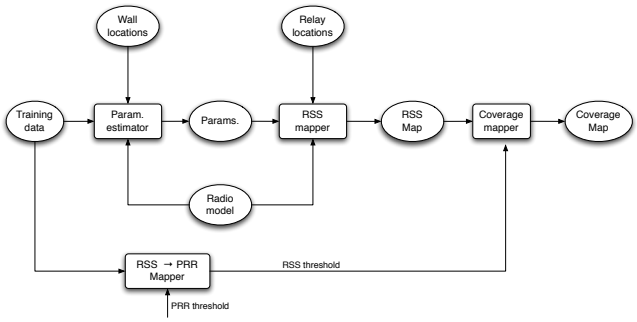


Figure 8. Radio Mapping Tool

approach allows to train the models and make predictions within minutes.

6 Empirical Evaluation of RMT

In this section, we analyze the performance of the RMT on data set previously collected for the empirical study in Section 4. We begin by assessing RMT's performance through a case study which highlights RMT's accuracy and that the intuitive nature of the radio maps RMT presents to the user. In the next subsection, we show that the automatic wall classification model which RMT uses by default achieves significantly better accuracy than either the basic log-normal model or the manual wall-classification model.

We characterize the accuracy of RMT's coverage predictions by its resulting false positive and false negative rates. In contrast to the previous section, the false positive and false negative rates discussed here refer to the prediction coverage rather than the RSS threshold. In this context, a false positive occurs when RMT predicts coverage where there is none; similarly, a false negative occurs when RMT predicts no coverage but ground truth data indicates otherwise.

6.1 Representative Examples

The case-study is designed to emulate the use of RMT to predict the coverage of one or two relays. In order to illustrate the efficacy of our automatic wall-classification model, we present results with the unmodified RMT as well as with a version of RMT that has been modified to use the basic log-normal model. To highlight RMT's accuracy when using only a small amount of training data, we randomly select 20% of our experimental data for training, while the remaining 80% is used as ground truth data for testing. The data is divided into training and testing sets through the same sampling strategy described in Section 4. For the purposes of this study, we define a "good link" to have a PRR higher than 80%. Using the RSS threshold selection technique previously discussed, an RSS threshold of -87 dBm was selected.

We first present the case when one relay (node 175 in Figure 9) is deployed. The results are summarized in Table 1.

Figure 9(a) plots the RSS predictions when the log-normal model is used. Since the log-normal model does not account for wall attenuation, the contour graph consists of concentric circles. RMT also plotted the -87 dBm line that delineates the relay's coverage area. Figure 9(b) shows the corresponding coverage map. The log-normal model predicts coverage correctly at only 59% of tested locations and had over 22 false positives as summarized in Table 1. The high false negative rate highlights that the log-normal model makes overly optimistic coverage predictions i.e., it predicts coverage where there is none.

Figure 9(c) presents the RSS predictions made by the automatic wall classification model. The results indicate that the relay's coverage is non-isotropic due to wall attenuation. A close examination shows significant signal attenuation caused by walls. This is especially apparent close to relay 187 (near the bottom-left corner) where the signal is significantly attenuated by the wall. Another place where the impact of walls is evident are the finger-like features present

	Log Normal	Automatic wall classification
Correct predictions	32	46
False positives	22	2
False negatives	0	6

Table 1. Accuracy of coverage predictions for one relay when using 20% of the data set for training

	Log Normal	Automatic wall classification
Correct predictions	39	46
False positives	15	1
False negatives	0	7

Table 2. Accuracy of coverage predictions for two relays when using 20% of the data set for training

in the predicted coverage area. These features are the result of changes in the number of walls through which the signals pass. This is most apparent when the signal propagate through the building's small hallway, as labeled near the center of the map.

Figure 9(d) shows the corresponding relay coverage. The automatic wall class model achieves a prediction accuracy of 85%. Compared to the log-normal model, the automatic wall class model reduces the number of false positives from 22 to 2. We note that the automatic wall classification model incurred a higher number of false negatives than the log-normal model. However, from a practical standpoint, false positives may have a more significant impact on the end-user. False positives represent locations where the model predicts coverage even though there is none, which could lead to coverage gaps. In contrast, a slight increase in the number of false positives may be acceptable since it would only lead to slightly denser network.

We also note that most of the false-positive and false-negative locations are close to the predicted coverage border. We expect that the coverage prediction could be improved by targeted sampling near the border. This highlights the use of RMT as an interactive tool to guide the user about where to collect additional coverage measurements.

We now consider a case where two relays are deployed at nearly opposite sides of the building, as shown in Figure 10 and summarized in Table 2. RMT combines the RSS predictions from each relay by taking the maximum RSS prediction at each point.

Figure 10(a) shows the RSS values predicted by the log-normal model. This prediction suggests that the two selected relays cover a large fraction of the floor plan. However, as may be observed in Figure 10(b), in reality a significant portion of the building is not covered with a PRR higher than 80%. The log-normal model achieves a prediction accuracy of 72%, incorrectly predicting coverage at 15 locations where there ground-truth data indicates a hole in coverage.

Figures 10(c) and 10(d) show the corresponding RSS and coverage predictions made using the automatic wall-classification model. It is worth noting that the model predicts a thin corridor near node 66 where there is no coverage. This highlights the fact that coverage regions may be disconnected when considering the attenuation of walls. The automatic wall classification model had only one false negative and seven false positives, resulting in a prediction accuracy of 85%.

6.2 Detailed Empirical Results

In this section, we compare in more detail the statistical performance of three different models: log-normal, manual wall-classification, and automatic wall-classification. Based on these results we provide several additional insights regarding the radio mapping problem.

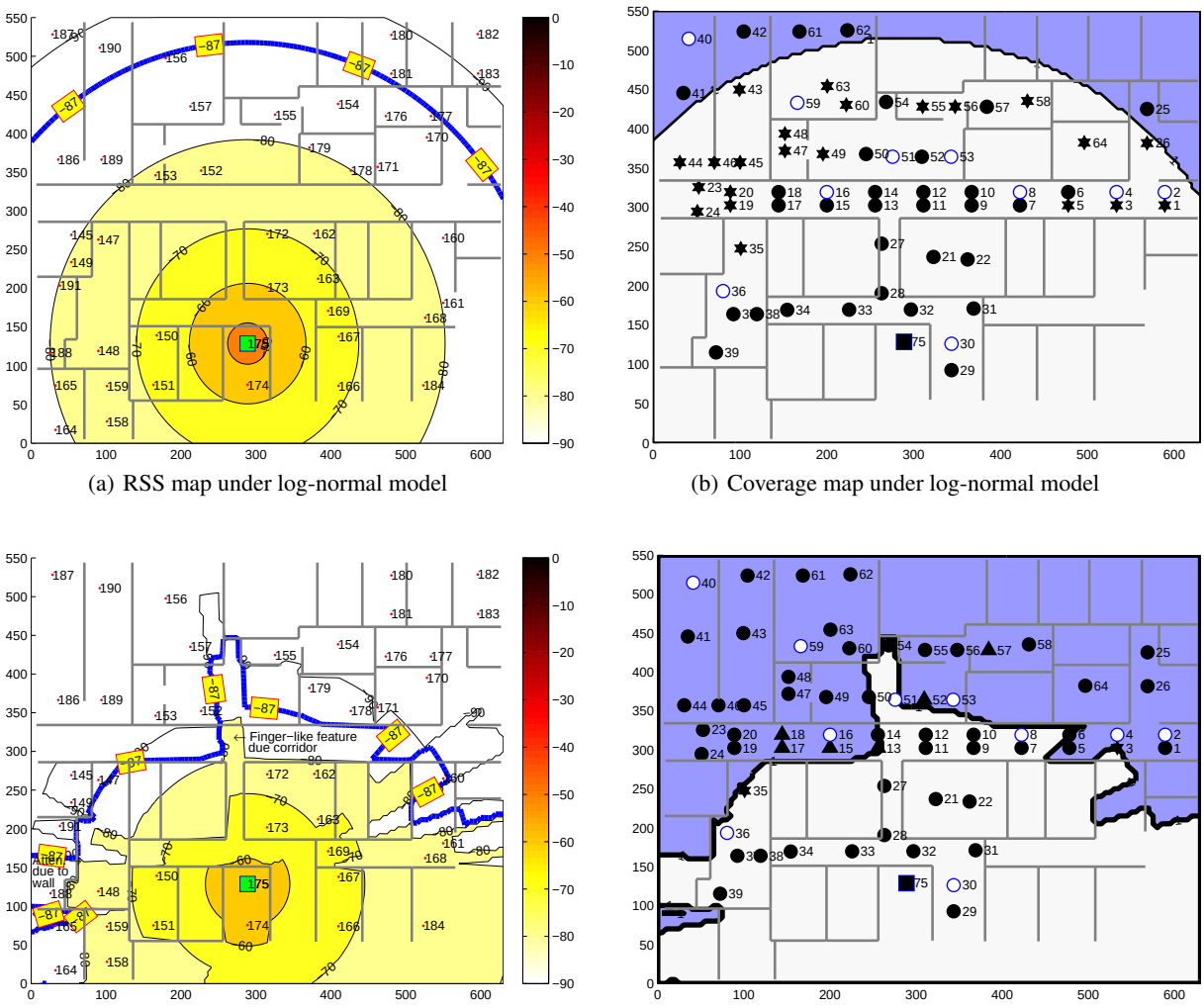
To understand how these models are affected by the amount of training data available to them, we performed a new set of experiments with the proportion of training data varied from 20% to 80% in increments of 10%. For each model at proportion, we randomly selected ten pairs of relays from the network and analyzed the false positive and false negative rates for the remaining test data. Like the previous experiment, we select a PRR threshold of 80%. As in Section 4, we present the average of 10 runs with error bars indicating the 90% confidence intervals, though the error bars are not always large enough to be clearly visible.

Figure 11(a) plots the false positive rates of these three models. As also seen in the case study, the log-normal model suffers from numerous false positives, with a false positive rate ranging from 51%–56%. In contrast, the models that incorporate wall information have significantly better false positive rates, reducing the false positive rate by 41% when using only 20% of the training data. At the 20% data point, the two wall classification models achieve comparable performance. For higher fractions of training data, the automatic wall classification achieves a false positive rate as much as 20% lower than the manual wall classification model. These results demonstrate that the proposed radio mapping technique with automatic wall classification achieves the lowest false positive rates, allowing the user to better identify coverage holes.

Figure 11(b) plots the false negative rates for the same training data. The log-normal model achieved the lowest false negative rate, of 4%. However, this comes at the cost of incurring a high false positive rates. The false positives for manual and automatic wall classification were 7% and 9%, respectively. These rates are comparable with the 10% false negative rate imposed in selecting the RSS threshold. We note that if the signal strength prediction was perfectly accurate, the the false positive rates in prediction coverage should equal 10%. As previously discussed, a moderate increase in false negative may be acceptable since it would only result in a slightly denser network.

Based on these results, we draw the following insights:

- Models that account for the impact of walls reduce the false positive rate by as much as 41%.
- The automatic wall classification model not only requires less user information, but also may reduce the false positive rate by as much as 20%.
- Accurate predictions of coverage are feasible even when the link quality to relays is sampled at only ten locations.



(a) RSS map under log-normal model

(b) Coverage map under log-normal model

(c) RSS map under automatic wall-classification model

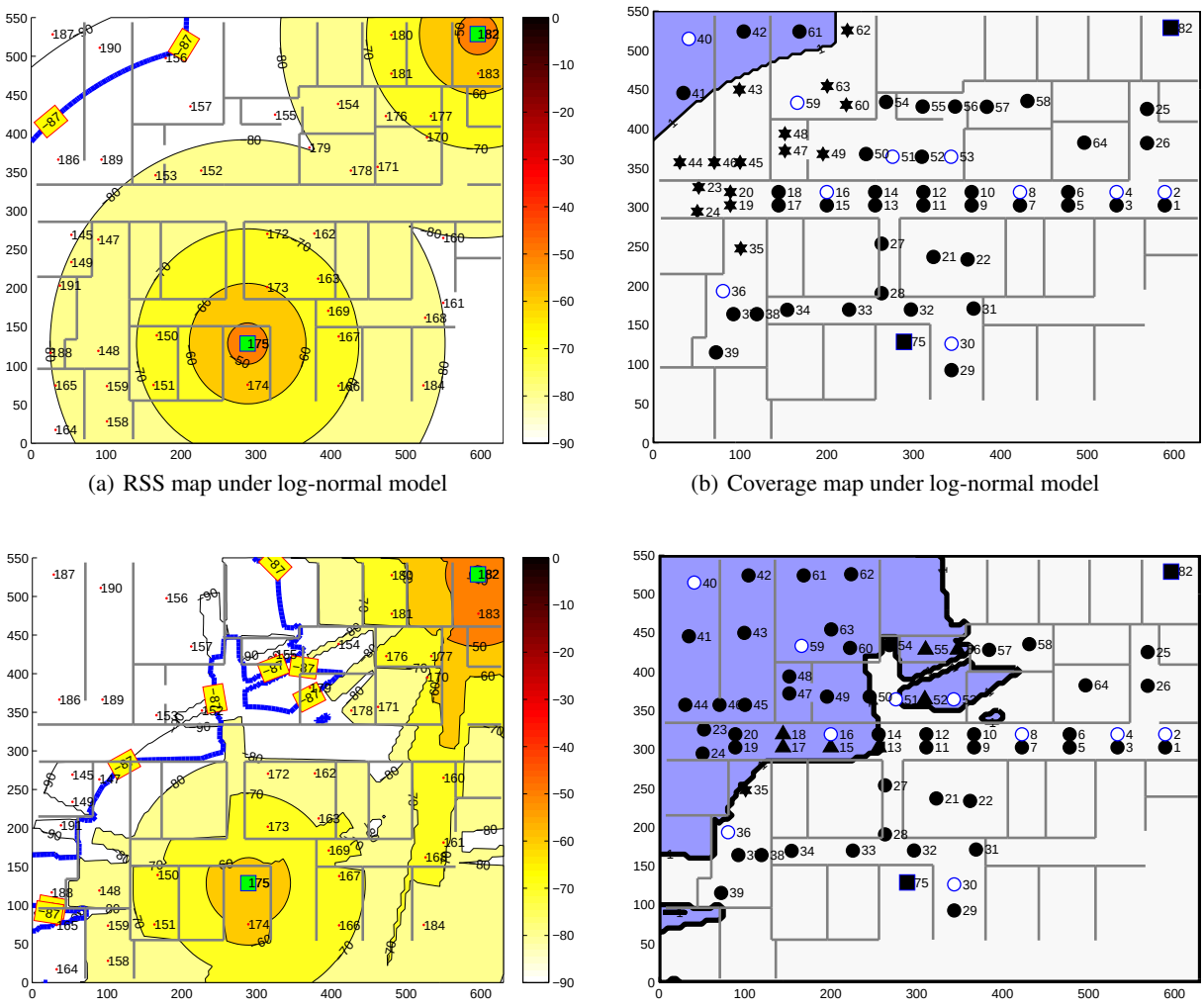
(d) Coverage map under automatic wall-classification model

Figure 9. Coverage and RSS predictions for one relay marked by a filled rectangle. The predicted relay coverage has a white background while the uncovered region is gray (blue on color printers). Different marker types are used to distinguish between training data (hollow circles), correct predictions (filled circles), false positives (filled stars), and false negatives (filled triangles).

7 Related Work

Modeling signal propagation is a challenging problem that has attracted a great deal of interest in the wireless communication community. Reviews of various radio propagation models for indoor and outdoor environments may be found in [1, 10, 12]. These models fall into two different categories: small-scale fading and large-scale path loss. The small-scale fading models use statistical techniques to describe variations in signal strength over a few wavelengths [19]. In contrast, large-scale path loss models predict average signal strength at distances significantly larger than the radio's wavelength. This paper focuses on the latter because large-scale path loss is more applicable to the radio mapping problem.

Most commonly, large-scale path loss models are derived by using empirical measurements to fit a small number of parameters (e.g., distance between sender and receiver, or walls) which significantly affect signal propagation. These models have the advantage of being intuitive to use and require minimal computational overhead. At the other end of the complexity spectrum, researchers have proposed site-specific techniques which involve ray tracing [16, 20]. [5] presents a tool for predicting signal strength of 802.11 access points at different locations through ray tracing techniques. A fundamental limitation of ray-tracing techniques is that they rely on the user to provide locations and attenuation coefficients for each partition or obstacle in the environment. Moreover, these techniques are usually computationally demanding and therefore not suitable for an interactive tool as we have built. We observe in Sections 4 and 6 that complex models are not necessarily better for radio mapping



(a) RSS map under log-normal model

(b) Coverage map under log-normal model

(c) RSS map under automatic wall-classification model

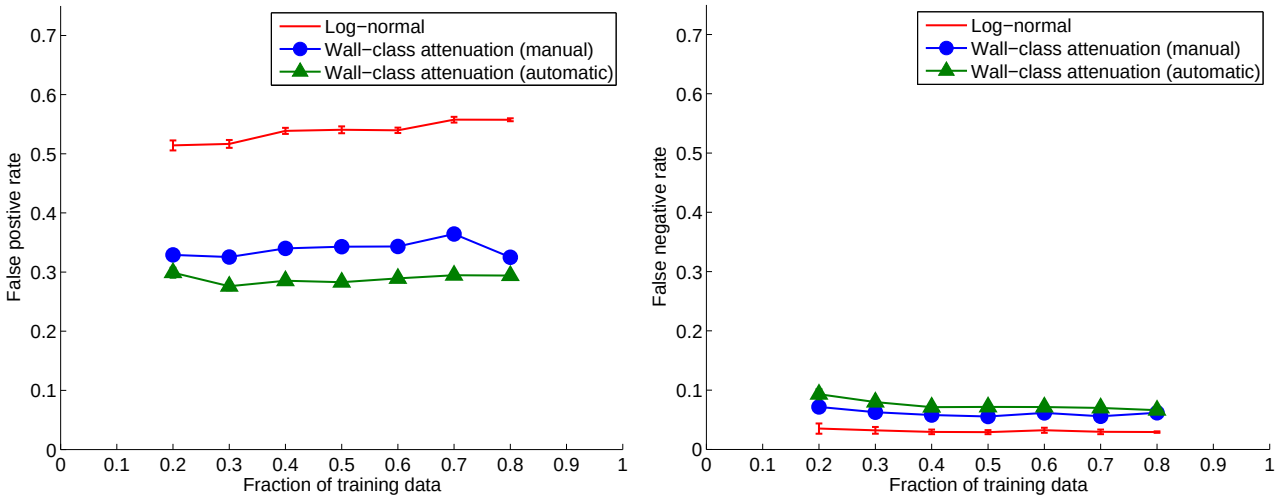
(d) Coverage map under automatic wall-classification model

Figure 10. Coverage and RSS predictions for two relays marked by filled rectangles

purposes, because there is a fundamental trade-off between model complexity and the amount of training data needed to accurately estimate the model's parameters.

[18] proposes a framework for assessing the signal strength and coverage of 802.11 mesh networks in outdoor environments. The paper proposes a radio propagation model which relies on the extraction of topological features from satellite images such as those provided by the Google Maps service. While our study and [18] both highlight the importance of topological information in assessing signal strength, [18] relies on a sectorization approach similar to the one considered in Section 2 which our own study finds to be ineffective. The different results owe to the different testing environments: in a complex indoor environment, wall attenuation dominates the effect of non-isomorphic radiation. Moreover, our study proposes a model based on wall classification which we expect would be inappropriate for an outdoor survey, since we do not expect different buildings to have similar attenuations.

A number of studies evaluate the properties of low-power wireless links [7, 15, 17, 22, 23, 29]. Our own empirical study focuses on predicting the signal strength at any location in space rather than characterizing existing links. Nevertheless, our empirical study evaluates the applicability of several insights made in these studies with regards to radio mapping. Prior studies have concluded that variations in hardware calibration may lead to significant variations in signal strength and packet reception rates [22, 22, 26]. We capture this effect by assigning a different α parameter to each node. [26, 29] demonstrate an angular dependency on signal strength caused by the non-isotropic radiation pattern of real-world antennas. We find in Section 4 that modeling this behavior is not effective for radio mapping purposes, because it introduces many more parameters that require large amounts of training data to adequately estimate. [15, 26, 30] use the log-normal model to predict link quality in simple indoor environments where nodes have line-of-sight. In contrast, our work targets environments where nodes do not always have



(a) False positive rate comparison

(b) False negative rate comparison

Figure 11. Comparison of mapping accuracy

direct line-of-sight; we show in Section 4 that this model has poor performance in such environments.

The sensor network community has recognized the importance of ensuring sensing or communication coverage early on. However, the majority of the work on coverage relies on geometric models that fundamentally assume that a node's communication/sensing range is circular [6, 9]. This assumption has been shown by many empirical studies to be incorrect, as does our case study in Section 6. More closely related to our work are two recent papers which look at sensing coverage. [14] proposes a framework which uses Gaussian processes to model sensing and communication costs. A disadvantage of Gaussian processes is that they cannot effectively model discontinuities such as those observed when a signal passes through walls. In contrast, the propagation model used by RMT explicitly models wall attenuation, which our study in Section 4 shows to be significant. [11] proposes a method for determining a sensor's sensing radio range through hierarchical sampling. This approach is complementary to our own, since it deals with efficient sampling strategies for refining coverage boundaries; our work focuses on processing the collected samples to predict coverage. Moreover, [11] exclusively addresses sensor coverage while our study deals with radio coverage.

8 Conclusion

Radio mapping is a challenging problem for real indoor environments due to signal attenuation through walls, complex signal propagation behavior, and the need to reduce the number of sampling measurements. This paper addresses this important challenge by developing a practical and effective radio mapping approach for indoor environments.

We first perform an in-depth empirical analysis of several signal propagation models in an office building. Our analysis shows the importance of balancing the accuracy of the model against the number of model parameters that need be estimated based on limited measurement. Our empirical results identify the wall-classification model as the most practical and effective model for indoor environments.

We then propose a practical algorithm to predict the RSS between different locations based on a small number of measurements. A key novelty of our algorithm lies in its ability to automatically classify walls into a small number of classes with different degrees of signal attenuation. Empirical results show that our automatic wall classification scheme results in more accurate RSS prediction than a manual classification based on architectural knowledge.

We have developed a practical Radio Mapping Tool to predict the radio coverage of relay placements. RMT has several salient features. (1) It requires minimal information about the indoor environment. The only knowledge about the environment that RMT needs are the wall locations, which may be extracted from existing floor plans. (2) RMT can accurately predict radio coverage based on a small number of measurements, which can significantly reduce the cost of network deployment and maintenance. (3) RMT features computationally efficient algorithms that allow users to quickly assess and adjust the coverage of a potential relay placement.

An empirical evaluation in the office building showed that RMT achieves as much as 41% fewer false positives compared to the log-normal model with a false negative rate of 9% based on sampling only 20% of the locations of interest. Our results demonstrate that RMT is a practical and accurate tool which can be used to facilitate the efficient deployment and robust operation of wireless sensor networks for indoor environments.

9 References

[1] J. Andersen, T. Rappaport, and S. Yoshida. Propagation measurements and models for wireless communications channels. *IEEE Communications Magazine*, 1995.

[2] R. Bernhardt. Macroscopic diversity in frequency reuse radio systems. *IEEE Journal on Selected Areas in Communications*, 1987.

[3] M. D'Souza, T. Wark, and M. Ros. Wireless localisation network for patient tracking. In *Proceedings of ISSNIP*, 2008.

[4] J. Farringdon, A. J. Moore, N. Tilbury, J. Church, and P. D. Biemond. Wearable sensor badge and sensor jacket for context awareness. In *Proceedings of ISWC*, 1999.

[5] S. J. Fortune, D. M. Gay, B. W. Kernighan, O. Landron, R. A. Valenzuela, and M. H. Wright. Wise design of indoor wireless systems: Practical computation and optimization. *IEEE Computational Science and Engineering*, 1995.

[6] S. Funke, A. Kesselman, F. Kuhn, Z. Lotker, and M. Segal. Improved approximation algorithms for connected sensor cover. *Wireless Networks*, 2007.

[7] 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, UCLA Computer Science Department, 2002.

[8] R. K. Ganti, P. Jayachandran, T. F. Abdelzaher, and J. A. Stankovic. Satire: a software architecture for smart attire. In *Proceedings of MobiSys*, 2006.

[9] H. Gupta, Z. Zhou, S. R. Das, and Q. Gu. Connected sensor cover: self-organization of sensor networks for efficient query execution. *IEEE/ACM Transactions Networking*, 2006.

[10] H. Hashemi. The indoor radio propagation channel. *Proceedings of the IEEE*, 1993.

[11] J. Hwang, T. He, and Y. Kim. Exploring in-situ sensing irregularity in wireless sensor networks. In *Proceedings of SenSys*, 2007.

[12] M. Iskander and Z. Yun. Propagation prediction models for wireless communication systems. *IEEE Transactions on Microwave Theory and Techniques*, 2002.

[13] J. G. Ko, T. Gao, , and A. Terzis. Empirical study of a medical sensor application in an urban emergency department. In *Proceeding of BodyNets*, 2009.

[14] A. Krause, C. Guestrin, A. Gupta, and J. Kleinberg. Near-optimal sensor placements: maximizing information while minimizing communication cost. In *Proceedings of IPSN*, 2006.

[15] 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 SenSys*, 2006.

[16] J. McKown and J. Hamilton, R.L. Ray tracing as a design tool for radio networks. *IEEE Network*, 1991.

[17] N. Reijers, G. Halkes, and K. Langendoen. Link layer measurements in sensor networks. In *IEEE MASS*, 2004.

[18] J. Robinson, R. Swaminathan, and E. W. Knightly. Assessment of urban-scale wireless networks with a small number of measurements. In *Proceedings of MobiCom*, 2008.

[19] A. Saleh and R. Valenzuela. A statistical model for indoor multipath propagation. *IEEE Journal on Selected Areas in Communications,*, 1987.

[20] K. Schaubach, N. Davis, and T. Rappaport. A ray tracing method for predicting path loss and delay spread in microcellular environments. In *Proceedings of IEEE Vehicular Technology Conference*, 1992.

[21] S. Seidel, T. Rappaport, M. Feuerstein, K. Blackard, and L. Grindstaff. The impact of surrounding buildings on propagation for wireless in-building personal communications system design. In *Proceedings of IEEE Vehicular Technology Conference*, 1992.

[22] D. Son, B. Krishnamachari, and J. Heidemann. Experimental study of the effects of transmission power control and blacklisting in wireless sensor networks. In *Proceedings of IEEE SECON*, 2004.

[23] K. Srinivasan and P. Levis. RSSI is under appreciated. In *Proceedings of EmNets*, 2006.

[24] K. Srinivasan and P. Levis. Rssi is under appreciated. In *Proceedings of EmNets*, 2006.

[25] J. A. Stankovic, Q. Cao, T. Doan, L. Fang, Z. He, R. Kiran, S. Lin, S. Son, R. Stoleru, and A. Wood. Wireless sensor networks for in-home healthcare: Potential and challenges. In *Proceeedings of HCMDSS*, 2005.

[26] T. Stoyanova, F. Kerasiotis, A. Prayati, and G. Papadopoulos. Evaluation of impact factors on rss accuracy for localization and tracking applications. In *Proceedings of MobiWac*, 2007.

[27] Texas Instruments. *2.4 GHz IEEE 802.15.4 / ZigBee-ready RF Transceiver*.

[28] A. Wood, J. Stankovic, G. Virone, L. Selavo, Z. He, Q. Cao, T. Doan, Y. Wu, L. Fang, and R. Stoleru. Context-aware wireless sensor networks for assisted living and residential monitoring. *IEEE Network*, 2008.

[29] G. Zhou, T. He, S. Krishnamurthy, and J. A. Stankovic. Models and solutions for radio irregularity in wireless sensor networks. *ACM Transactions on Sensor Networks*, 2006.

[30] M. Zuniga and K. Bhaskar. An analysis of unreliability and asymmetry in low-power wireless links. *ACM Transactions on Sensor Networks*, 2007.