A Hidden Environment Model for Constructing Indoor Radio Maps

Zhe Xiang* Hangjin Zhang+ Jian Huang* Song Song* Kevin C. Almeroth+

*IBM China Research Lab
Beijing, China
{xiangzhe,jianh,soong}@cn.ibm.com

+Department of Computer Science
University of California–Santa Barbara
{hangjin,almeroth}@cs.ucsb.edu

Abstract

Constructing indoor radio maps plays an important role in many services and applications such as wireless base station planning. In this paper, we propose a hybrid approach to constructing indoor radio maps by developing a novel indoor signal propagation model, called the Hidden Environment Model (HEM). The model is a hybrid because it combines on-site measurements with a number of different types of calculations. As part of this model, we introduce the idea of an Environment Factor Matrix (EFM). The EFM represents a model of the environmental features that affect radio attenuation. We also develop a Lazy Sampling Algorithm to help generate the EFM. The goal of the Lazy Sampling Algorithm is to balance the number of measurements that need to be taken with our model's accuracy. The goal is to minimize the measurement workload while maintaining satisfactory accuracy. We evaluate our model by comparing the radio maps calculated from the model to a radio map obtained by exhaustive measurements. The results show that our Hidden Environment Model achieves good accuracy.

1 Introduction

With the development of pervasive wireless communication systems, the analysis and construction of indoor radio models has become increasingly important. Knowing the signal strength in an indoor wireless network is critical for a number of wireless services and applications, such as indoor base station planning and wireless security. Moreover, indoor wireless positioning systems rely heavily on the accuracy of indoor radio propagation maps [1, 10, 5, 2]. To meet the needs of these various applications, efficient construction of accurate radio maps becomes an important issue.

Traditionally, indoor radio map construction falls into two categories: on-site measurement [1, 10] and calculation from theoretical models [8, 3, 4, 7, 6]. On the one hand, on-site measurement techniques are quite time-consuming and cumbersome to generate. On the other hand, due to the severe multi-path effects, dead-spots, noise and interference in an indoor environment, the theoretical model-based approaches tend to be rather inaccurate. Very few of these techniques achieve a good balance between accuracy and efficiency. They either have limited accuracy or are difficult to build in practice.

In this paper, we introduce a *hybrid* radio map construction approach that combines indoor radio propagation model with a *small number* of on-site signal measurements. Our goal is to develop a radio map construction method that is both accurate and efficient (i.e., easy to generate). In our proposed *Hidden Environment Model (HEM)*, we build an *Environment Factor Matrix (EFM)* which models the distinct effects of the indoor environment on signal attenuation. It is constructed using on-site measurements combined with theoretical interpolation and approximation. A *Lazy Sampling Algorithm* is also proposed to optimally select the set of on-site measurements to be taken. The *Lazy Sampling Algorithm* effectively reduces the measurement workload while ensuring model accuracy.

The *Hidden Environment Model* exhibits the advantages of accuracy, efficiency and flexibility. As part of our evaluation of the model, we carried out a series of experiments at the IBM China Research Lab using 2.4 GHz 802.11b radios. Our results show that the mean square error of the radio maps generated from our model is less then 6 dBm.

The remainder of this paper is organized as follows. Related work is described in Section 2. In Section 3, the *Hidden Environment Model* is introduced and described in detail. In Section 4, we conduct a series of experiments to validate the model and evaluate its performance. The paper is concluded in Section 5.

2 Related Work

There are two approaches to construct indoor radio maps. Most recent efforts are focused on using model-based methods[1, 8, 3]. The work in this area can further be divided into two sub-categories: (1) free space signal propagation models; and (2) indoor signal propagation models.

In free space environments, signal strength is related only to the distance between the transmitter and receiver. When dBm is chosen as the measurement metric, the signal propagation model can be formulated as:

$$P(d)[dBm] = P(d_0) - 10\alpha log(d/d_0). \quad (1)$$

In Equation (1), $P(d)$ is the signal power received at a distance d, $P(d_0)$ is the power at a reference distance d_0, and α is the path loss exponent.

For an indoor environment, the wireless channel is noisy and the radio signal propagation is dominated by reflections, diffraction and scattering caused by the structures and obstacles within the building. These factors make the signal strength a complex function of distance and environment. A widely used and accepted empirical model for indoor environments is the Floor Attenuation Factor (FAF) model described by Seidel and Rappaport [8]. The FAF model include a Wall Attenuation Factor (WAF) to consider the effects of obstacles (i.e. walls). The FAF model is described by:

$$P(d)[dBm] = P(d_0) - 10\alpha log(d/d_0) - W \times WAF, \quad (2)$$

where W is the number of obstructions between the transmitter and the receiver, and WAF is the wall attenuation factor. The FAF model requires knowledge about the basic layout of the environment. In addition, the WAF is largely dependent on construction materials. One of the most significant weaknesses of the FAF model is that it does not take into account a number of important effects, including signal reflection, diffraction and scattering.

More recent work develops radio maps using a two-step parametric and measurement-driven ray-tracing approach to account for absorption and reflection characteristics of various obstacles [11]. Again, their method requires precise knowledge about the building layout and structure.

In summary, most earlier work in this area either requires heavy human workload, or requires difficult-to-obtain model parameters. There is a real need for a model that is both more accurate and easier to build. This is the goal of our proposed solution.

3 The Hidden Environment Model

In this section, we describe our proposed *Hidden Environment Model*. The core components of the Hidden Environment Model are a model equation and an *Environment Factor Matrix (EFM)*. The EFM is constructed based on a set of *Group Measurements*. In turn, these Group Measurements are determined three methods: Measurement, Interpolation and Approximation. The final component, the *Lazy Sampling Algorithm,* is designed to optimally select the right set the *Group Measurements*. The Lazy Sampling Algorithm is designed to achieve a balance between measurement workload and model accuracy. Each of these components will be described in detail in this section.

3.1 Model Equation

One of the main ideas behind the *Hidden Environment Model* is to separate the calculation of radio signal propagation from the effects of the environment. Rather than to develop a complex mathematical function to represent every environmental effect, we determine the effects using on-site measurements and then input this information into the model.

The formula used in our model is as follows. Assuming a base station is placed at Point A and the receiver at Point B, the signal power in our Hidden Environment Model is formulated as:

$$P_A(B)[dBm] = P_A(C) - 10\alpha log(|AB|/|AC|) - Env(A, B). \quad (3)$$

Here, $P_A(B)$ represents the signal strength received at Point B; and $P_A(C)$ is the signal received at reference Point C. The constant, α, is the path loss exponent, $|AB|$ and $|AC|$ represent the distance between Points A and B and A and C respectively. $Env(A, B)$ is the environment factor matrix modeling the effects of the surrounding environment on the signal attenuation between Points A and B.

The principle behind Equation (3) is that two factors contribute to overall signal attenuation. The first factor is signal loss due to the distance between the transmitter and the receiver. The other factor is environmental effects, captured by the term $Env(A, B)$. This equation can be considered as a revised version of the free-space model equation. In addition to free-space path loss, we account for the effects of the in-door environment with $Env(A, B)$.

If we fix the distance between the reference point and the base station point and let $|AC| = 1$, Equation (3) is simplified to:

$$P_A(B)[dBm] = C - K \times log(|AB|) - Env(A, B). \quad (4)$$

$C = P_A(C)$ is determined solely by the physical specifications of the base station (e.g. transmission power). In theory, $K = 10\alpha$ is a constant (about 20). But in practice, we found it is slightly dependent on the environment. In our experiments, we obtain a value for K by averaging a set of measurements.

The above equations are designed for omni-directional antennas. However, if the transmitter antenna is directional, (e.g., a typical Yagi antenna [9]) the equation is modified to:

$$P_A(B)[dBm] = C - (K_1 + \theta K_2) \times log(|AB|) - Env(A, B), \quad (5)$$

where θ is the angle between the Line $|AB|$ and the main axis of the directional antenna.

3.2 The Environment Factor Matrix

Suppose there are N points of interests, $P_1, P_2, \ldots, P_N$, in a given environment, we define the *Environment Factor Matrix* E as:

$$E_{N \times N} = [e_{i,j}], \quad (6)$$

where $e_{i,j} = Env(P_i, P_j)$. This matrix therefore represents the environmental effects at every pair of Points (P_i, P_j) in the model area.

In order to build the Environment Factor Matrix for a given environment, we use a combination of three methods: Measurement, Interpolation and Approximation. Each of these is described below.

Measurement: If we change the format of equation (4) to the following:

$$Env(P_i, P_j) = C - K \times log(|P_iP_j|) - P_i(j), \quad (7)$$

we see that the Environment Factor $Env(P_i, P_j)$ can be easily obtained by measuring the Signal Strength $P_i(j)$. Recall that $P_i(j)$ represents the signal strength measured at Point P_j when the base station is placed at Point P_i.

Clearly, direct measurement is the most accurate method for determining the Environment Factor Matrix. However, this method requires placing the base station at either P_i or P_j and taking a measurement at the other point. This requirement requires a large number of measurements. Therefore, two mathematical methods, Interpolation and Approximation, are introduced to lessen the workload.

Interpolation: To apply interpolation, we first assume that environment factors are symmetric, i.e., $e_{i,j} = e_{j,i}$ We also assume that environment factors obey the rule of linear addition and subtraction. As shown in Figure 1, if Points P_i, P_j, P_k are collinear, we assume that $e_{j,k} = e_{i,j} \pm e_{i,k}$. The plus or minus sign depends on the location relationship of the three points. When Points P_j and P_k are on different sides of P_i, the two terms are added together; otherwise, the terms are subtracted. While the Interpolation method is only an estimation, experiment results show it to be reasonable. The Interpolation method does not require the base station to be an endpoint, but it does require the base station to be on the same line. The next method, Approximation, relaxes this requirement but trades off accuracy.

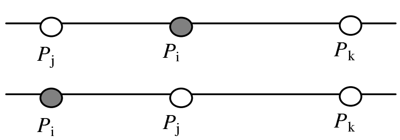

Figure 1. Interpolation

Approximation: As shown in Figure 2, if three Points P_i, P_j, P_k, are not collinear, but the distance between Point P_i and Line $|P_jP_k|$ is less than some Threshold, ϵ, we still apply the linear addition rule to approximately calculate $e_{j,k}$, i.e., $e_{j,k} \approx e_{i,j} \pm e_{i,k}$. Although the Approximation method may introduce more error than the Interpolation method, the tradeoff is that it greatly reduces the measurements workload.

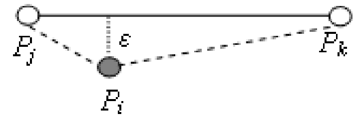

Figure 2. Approximation

We have now described three methods for constructing the *Environment Factor Matrix*. All of them require some preliminary set of measurements, namely $P_i(j)$ and $P_i(k)$. The next challenge is then to determine the set of P_i's. For this, we design the *Lazy Sampling Algorithm*. Before describing this algorithm though, we first describe how on-site *Group Measurements* are taken.

3.3 Group Measurements

Recall that in order to obtain $P_i(j)$, we need to place the base station at position P_i; place the receiver at position P_j; and measure the signal strength. Because it is likely to be much easier to move the receiver than to move the base station, we conduct one round of *Group Measurements* by placing the base station at a fixed point (called the Transmitter Point), and then move the receiver to each of the other $N - 1$ points to measure the signal strength.

Clearly, the overall measurement workload is determined by the number of rounds of *Group Measurements* that need to be taken. Determining the right set of Transmitter Points is critical. Therefore, we have developed the *Lazy Sampling Algorithm* to optimize Transmitter Point selection.

3.4 The Lazy Sampling Algorithm

The purpose of the *Lazy Sampling Algorithm* is to optimally select a valid set of *Transmitter Points* where *Group Measurements* need to be taken. A good set of Transmitter Points should have the following property: for every pair of Points (P_i, P_j) where $1 \leq i, j \leq N$, the Environment Factor $e_{i,j}$ can be calculated from the Group Measurements taken at the Transmitter Points, using any of the three methods described earlier. This property means that for every pair of Points (P_i, P_j), there exists a Transmitter Point, P_t, such that the distance between P_t and the line $|P_iP_j|$ is less than ϵ.

Unfortunately, the selection of Transmitter Points is NP-Complete, because this problem can be reduced from another known NP-Complete problem, the Minimal Dominating Set problem. A graph $G = (V + P, E)$ can be constructed, where $V = \{P_i | 1 \leq i \leq N\}$ is the points set, $P = \{(P_i, P_j) | 1 \leq i, j \leq N\})$ is the pairs set, and $E = \{(v, p) | v \in V, p \in P, distance(v, p) \leq \epsilon\}$ defines the edges set in the graph. Finding the minimal set of Transmitter Points is equivalent to finding the minimal dominating set in the Graph G.

Our proposed *Lazy Sampling Algorithm* is a greedy algorithm. It first computes a *Coverage Set* for each point. A Coverage Set for Point P_i is defined as $CS(P_i) = \{(P_j, P_k) | distance(P_i, |P_j P_k| \leq \epsilon)\}$. At each stage, the point with the largest Coverage Set size is selected. Then this point and all nearby ($\leq \epsilon$) pairs are deleted from the Graph G. This step is repeated until all pairs are deleted from the graph. The principle of the *Lazy Sampling Algorithm* is straightforward. Intuitively, the point with the largest Coverage Set seems to be have more benefit and thus is selected first. Although the greedy algorithm is not optimal, in almost all cases the algorithm identifies a reasonable set of Transmitter Points.

4 Evaluation

In this section, we perform a series of experiments to both qualitatively and quantitatively study the accuracy of our model. Our goal is to show the accuracy of the model, regardless of which type of the transmitter antenna is used. In addition, we also quantify the tradeoff between accuracy and the number of Group Measurements.

We begin by describing our experimental setup, followed by a description of how the maps were constructed, and finally, the results.

4.1 Experiment Setup

The experiments reported in this paper were conducted in an area of the IBM China Research Lab. The area covers 240 m^2 and contains two rooms. We intentionally selected a non-rectangular area that contains different types of wall materials to demonstrate the accuracy of our model even in a complex environment. The layout of the area is shown in Figure 3. The area is uniformly divided into a 46×28 grid of cells. Each cell is $0.6m$ by $0.6m$. Among the $1,288(46 \times 28)$ cells, there are about 700 effective points (within the rooms).

The parameters and constants used in our evaluation are shown in Table 1. C and K are the parameters for an omni-directional antenna. C, K_1 and K_2 are used for a directional antenna. These values were obtained by averaging a small set of measurements. We also define the Distance Threshold ϵ as $1.2m$, a value that is two times the length of a cell.

We show the accuracy of our model by comparing the radio maps created by it to radio maps obtained through exhaustive on-site measurements. The similarity between two

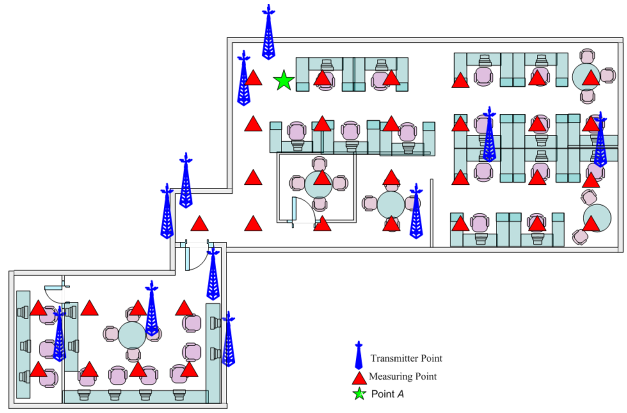

Figure 3. Layout of the Experiment Area

Parameter	Value
C	-45.0
K	-15.0
K_1	-6.0
K_2	-8.0
$\epsilon(m)$	1.2

Table 1. Parameters Used in Experiments

radio maps serves as the metric for accuracy. We use two metrics to evaluate the quality of the model-calculated radio map: the signal contour map and the numeric signal vector. The contour maps provide a visual (qualitative) means to compare the two maps, while the numeric signal vectors are used to compute the Mean Square Error (quantitative).

4.2 Calculating the Radio Maps

When applying the *Lazy Sampling Algorithm* on the test area, it outputs 11 Transmitter Points (the transmitter towers in Figure 3). We then conducted 11 rounds of simplified Group Measurements. The Group Measurements are simplified because we did not measure signal strength at all 700 effective points in the rooms, rather, we only measured signal strength at 33 uniformly distributed points (the small triangles shown in Figure 3). The signal strength at the other 660+ points was calculated using interpolation. After the Group Measurements were completed, we are able to calculate the Environment Factor Matrix and create the final the Hidden Environment Model.

Next we report on a comparison of the signal strength contours and then discuss the analysis of the numeric signal strength vectors.

4.3 Experiment Results and Analysis

In our experiments, we randomly selected a place to mount the base station (Point A shown in Figure 3), and

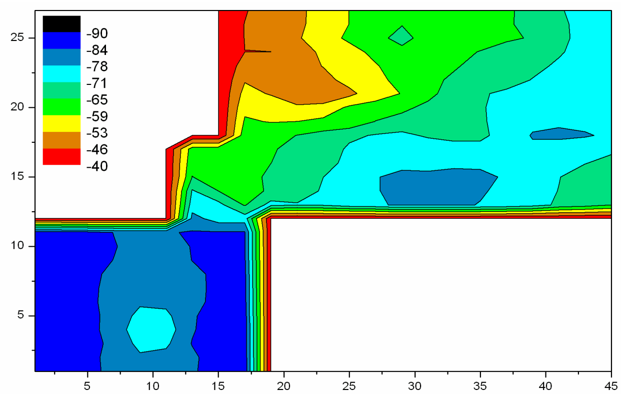

Figure 4. Measurement-Based Omni Radio Map

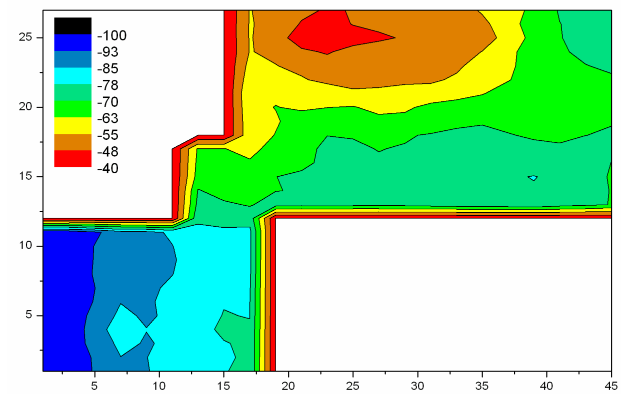

Figure 6. Measurement-Based Directional Radio Map

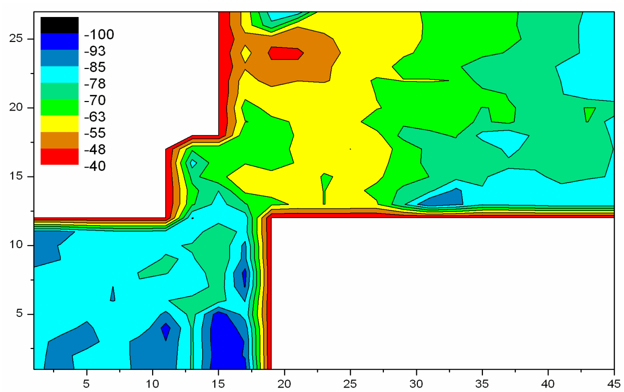

Figure 5. Model-Based Omni Radio Map

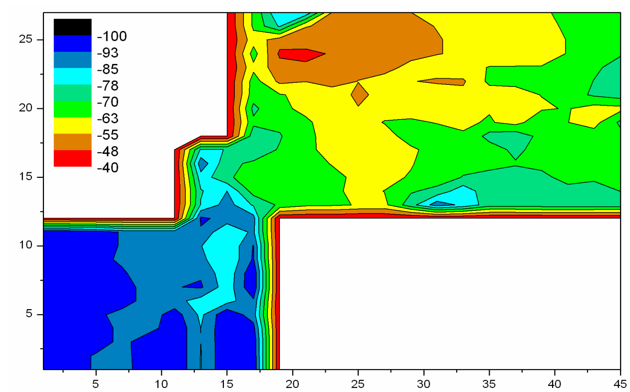

Figure 7. Model-Based Directional Radio Map

used the model to calculate the radio map it generates. In addition to evaluating the results for an omni-directional antenna, we also evaluated results for a directional antenna. We first describe an experiment using an omni-directional antenna and then one for a directional-antenna.

Figures 4 and 5 illustrate the results when an omni-directional antenna is used. Figure 4 plots the contour map from the real measured signal strength while Figure 5 plots the contour from our Hidden Environment Model. The two contours show a significant amount of similarity. The calculated radio map effectively captures the environmental effects on signal propagation. On the one hand, the signal is gracefully attenuated from the Transmitter Point to the right side of the main room since there are no significant obstacles in that direction. On the other hand, the signal is essentially blocked from passing through the concrete walls to the adjoining room. The visual similarity between the two maps helps to validate, at least qualitatively, the accuracy of our Hidden Environment Model.

We also conducted several experiments for directional antennas. In one set, the antenna was facing east (right), and in the other, the antenna was facing south (down). Due to space limitations, we only present the set of results with the antenna facing right. The two contour maps are shown in Figures 6 and 7.

As shown in Figures 6 and 7, the radio maps for the directional antenna are also qualitatively similar. Even though the Group Measurements are performed for an omni-directional antenna, the Environment Factor Matrix is constructed to be independent of the type of antenna. This result shows the flexibility of the model: the Environment Factor Matrix generated from an omni-directional antenna is also applicable to directional antennas.

For a quantitative comparison, we calculated the Mean Square Error (MSE) of the model-based radio map compared to the measurement-based radio map. The results for the three sets of experiments described above are shown in Table 2. Note that in all experiments, the error is within 6 dBm which effectively validates the accuracy of the *Hidden Environment Model*.

Experiment	MSE(dBm)
Omni Antenna	5.64
Directional Antenna (East)	5.70
Directional Antenna (South)	6.01

Table 2. Mean Square Error of Calculated Radio Map

4.4 Impact of Measurement Rounds on Accuracy

Given that reducing the measurement workload is one of the key goals of our Hidden Environment Model, we study how the number of Group Measurements affects the model's accuracy. Our analysis is based on varying the number of Group Measurements used to construct the Environment Factor Matrix. As we vary the number of measurements used, the Distance Threshold, ϵ, may not strictly hold, i.e., in some cases when only a few measurements are taken we may be forced to estimate values that are greater than ϵ meters away. The mean square error results for each set of experiments are plotted in Figure 8.

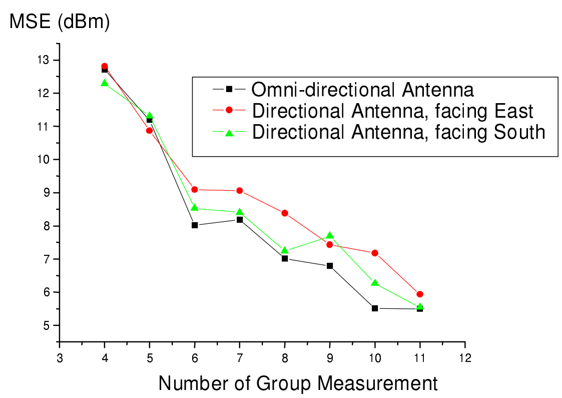

Figure 8. Impact of the Number of Group Measurements on Radio Map Accuracy

Figure 8 first shows that no matter what type of antenna is chosen, the radio map accuracy is essentially the same. Second, Figure 8 shows that as more measurements are taken, the mean square error decreases. When all 11 rounds of Group Measurements are taken, the error is less than 6 dBm. When only 4 rounds are taken, the error approaches 13 dBm. These results indicate that on-site measurements are a very important part of the model, and the set of Transmitter Points generated from the Lazy Sampling Algorithm are indispensable in maintaining accuracy.

5 Conclusions

Indoor radio maps play an important role in many services and applications, such as wireless positioning services and wireless base station planning. The traditional methods of radio map construction, including on-site measurement and model-based mathematical calculation, suffer from low accuracy, heavy workload and/or parameters that are hard to obtain.

In this paper, we proposed a hybrid approach to building indoor radio maps by combining an innovative indoor propagation model, called the *Hidden Environment Model*, and on-site measurement. By introducing the concept of an *Environment Factor Matrix*, the features of how an environment affects radio signal attenuation can be well modeled. Finally, we have developed a *Lazy Sampling Algorithm* to help design an optimal *Group Measurements* plan. The Lazy Sampling Algorithm allows a balance to be achieved between measurement workload and model accuracy.

Our Hidden Environment Model is useful when a radio map for arbitrary base station placement is needed. Currently we are working on applying our Hidden Environment Model to find the optimal base station placement for a radio positioning system.

References

[1] P. Bahl and V. Padmanabhan. Radar: An in-building rf-based user location and tracking system. In *Proceedings of IEEE Infocom*, pages 775–784, Tel-Aviv, Israel, March 2000.

[2] R. Battiti, T. Nhat, and A. Villani. Location-aware computing: a neural network model for determining location in wireless lans. In *Technical Report DIT-02-0083, University of TRENTO*, 2002.

[3] Y. Chen and H. Kobayashi. Signal strength based indoor geolocation. In *Proceedings of the IEEE International Conference on Communications*, pages 436–439, April 2002.

[4] H. Hashemi. The indoor radio propagation channel. In *Proceedings of the IEEE*, pages 943–968, July 1993.

[5] P. Myllymaki, T. Roos, H. Tirri, P. Misikangas, and J. Sievanen. A probabilistic approach to wlan user location estimation. In *Proceedings of the Third IEEE Workshop on Wireless LANs*, 2001.

[6] L. Rice. Radio transmission into building at 35 and 150 mc. In *Bell Systems Technical Journal*, pages 197–210, January 1959.

[7] S. Rice. Mathematical analysis of random noise. In *Bell Systems Technical Journal*, 1944.

[8] S. Seidel and T. Rapport. 914 MHz path loss prediction model for indoor wireless communications in multi-floored buildings. In *IEEE Transaction on Antennas and Propagation*, February 1992.

[9] H. Trzaska. Calibration of directional antennas and limitations in their use. In *IEEE Transaction on Instrumentation and measurement*, pages 1112–1116, January 2000.

[10] Z. Xiang, S. Song, J. Chen, H. Wang, J. Huang, and X. Gao. A wireless lan-based indoor positioning technology. In *IBM Journal on Research and Development*, pages 617–626, September 2004.

[11] G. Zaruba, M. Huber, F. Kamangar, and I. Chlamtac. Monte carlo sampling based in-home location tracking with minimal rf infrastructure requirements. In *Proceedings of the IEEE Globecom*, pages 3624–3629, December 2004.