User Location and Tracking in an In-Building Radio Network

Paramvir Bahl
Venkata N. Padmanabhan

February 1999

Technical Report
MSR-TR-99-12

Microsoft Research
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052

User Location and Tracking in an In-Building Radio Network[1]

Paramvir Bahl and Venkata N. Padmanabhan
Microsoft Research
{bahl, padmanab}@microsoft.com

Abstract

The proliferation of mobile computing devices and local-area wireless networks has fostered a growing interest in location-aware systems and services. A key requirement for enabling such services is user location and tracking. In this paper we address this problem in the context of a popular radio-frequency wireless network technology. Our approach is based on recording and processing real-time signal strength information available at multiple base stations positioned to provide overlapping coverage in the area of interest. We employ techniques that combine empirical measurements with signal propagation modeling. We present concrete experimental results that demonstrate the feasibility of estimating user location with a high degree of accuracy.

1 Introduction

The proliferation of mobile computing devices and local-area wireless networks has fostered a growing interest in location-aware systems and services. A key distinguishing feature of such systems is that the application information and/or interface presented to the user is, in general, a function of his or her physical location. The granularity of location information needed could vary from one application to another. For example, locating a nearby printer requires fairly coarse-grained location information whereas locating a book in a library would require fine-grained information.

While much research has focussed on developing services architectures for location-aware systems, less attention has been paid to the fundamental and challenging problem of user location and tracking, especially in in-building environments. The few efforts that have addressed this problem have typically done so in the context of infrared (IR) wireless networks. The limited range of IR networks, which facilitates user location, is a handicap for providing ubiquitous coverage in a building. Also, often the IR network is deployed for the sole purpose of locating people and does not provide traditional data networking services. For these reason we focus on radio-frequency (RF) wireless networks in our research. Our goal is to complement the data networking capabilities provided by RF networks with accurate user location and tracking capabilities, thereby enhancing the value of such networks.

Location determination in RF network requires analyzing signal characteristics. However, a major challenge arises due to the significant fluctuations in channel characteristics, both in time and space, typical in RF networks.

Our approach to the user location problem is as follows. We use signal strength information gathered at multiple receiver locations to *triangulate* the user's coordinates. To do the triangulation, we use both empirically determined and theoretically computed signal strength information.

Our experimental results are very encouraging. With high probability we are able to estimate a user's location to within a few meters of his/her actual location. This suggests that it is feasible for fine-grained location-aware services to be built on top of RF local-area wireless networks.

The remainder of this paper is organized as follows. In Section 2, we survey related work in location determination technologies. In Section 3, we discuss our research methodology. Section 4 contains the core of the paper where we present and analyze the empirical and the signal propagation modeling methods. A discussion of some related issues appears in Section 5. Finally, we present our conclusions in Section 6.

2 Related Work

Related work in the area of user location and tracking has been done in the following broad contexts: (1) in-

[1] This is an abridged version of a more detailed paper in preparation. Please contact the authors for additional information.

building IR networks, (2) wide-area cellular networks (based on RF), and (3) Global Positioning System (GPS).

The *Active Badge* system [Har94] was an early, significant contribution to the field of location-aware systems. In this system, a badge worn by a person emits a unique IR signal every 10 seconds. Sensors are placed at fixed positions within a building and as they receive the unique identifiers, the location manager software is able provide information about the person's location to the requesting services and applications. While the performance of this system is quite good, a major drawback is that the range of the IR system is fairly small, and consequently the building has to be wired up with a significant number of sensors. In the few places where such systems have been deployed, sensors have been physically wired in every room of the building. Such a system scales poorly, and incurs significant installation, configuration and maintenance cost. Also, IR tends to perform poorly in the presence of direct sunlight, which is common in rooms with windows.

Another system that is based on IR technology is described in [Azu93]. This system requires IR transmitters to be located at fixed positions inside the ceiling of the building. An optical sensor sitting on a head mounted unit senses the IR beacons and system software determines the position of the person. This system suffers from similar drawbacks as the Active Badge system.

The system described in [ATC97] uses pulsed DC magnetic fields to determine the position and orientation of the person with a high degree of precision. Multiple sensors are placed on body-mounted peripherals, such as data gloves and the output from the sensors is processed to determine location and orientation. This technology is used extensively in the computer animation industry. It is, however, quite expensive, and like IR, is severely range limited. Therefore, it is not suitable for wide-scale deployment.

Recently several location systems have been proposed for wide-area cellular systems [Tek98]. The technological alternatives for locating cellular telephones involve measuring signal attenuation, angle of arrival (AOA), and/or time difference of arrival (TDOA). Based on initial studies, the AOA and TDOA based system have been found to be promising.

A common variant of the angle-of-arrival technique is known as small aperture direction finding, which requires a complex antenna array at each of the cell site locations. The antenna arrays can in principle work together to determine the angle (relative to the cell site) from which the cellular signal originated. When several cell sites can determine their respective angles of arrival, the cell phone location can be estimated from the intersection of projected lines drawn out from the cell site at the angle corresponding to the signal's origin. Due to the nature of indoor environments, the angle of arrival technology cannot be used reliably to compute the position of a mobile user. Also, the TDOA technology requires very fine grained time synchronization between the transmitter and receiver, which is not feasible with current off-the-shelf mobile devices.

Some systems based on the Global Positioning Systems (GPS) have also been proposed [GPS99]. Unfortunately GPS transmissions are blocked by buildings, so the system does not operate indoors.

The Daedalus project [Hod97] developed a system that provides coarse level location tracking services. Beacon signals transmitted by the base station are augmented with location information. The mobile host estimates its location to be the same as that of the base station with the strongest signal. Given the relatively large size of cells, the location information is not likely to be very accurate.

Our work differs from previous work in that we tackle the problem of location determination and tracking on a widely available *radio frequency* based wireless network in an *in-building* environment. RF networks offer a significant advantage over IR networks in terms of range, scalability, deployment, and maintenance. With speeds up to 10 Mbps these systems have gained rapid acceptance and are being widely deployed in companies, schools, homes etc.

3 Research Methodology

In this section, we describe our research methodology. We begin with a description of our experimental testbed. We then discuss the data collection process, including tools we developed for this purpose. Finally, we describe the processing we performed on the data as a precursor to the analysis described in Section 4.

3.1 Experimental Testbed

Our experimental testbed is located on one floor of a 3-storey building. The layout of the floor is shown in Figure 1. The floor has dimension of 43.5 m by 22.5 m, an area of 980 sq. m (10500 sq. ft.), and includes more than 50 rooms.

We placed three base stations, BS_1, BS_2 and BS_3, at the locations indicated in Figure 1. Each base station was a Pentium-based PC running FreeBSD 3.0 quipped with a wireless adapter. Our mobile host, which was carried by the user being tracked, was a Pentium-based laptop computer running Microsoft Windows 95.

Each base station and the mobile host was equipped with a Digital RoamAbout™ network interface card (NIC), based on Lucent's popular WaveLAN™ RF LAN technology. The network operates in the 2.4 GHz license-free ISM (Industrial, Scientific and Medical) band. It has a raw data rate of 2 Mbps and a one-way delay of 1-2 ms. The range of the network, as specified in [Roa96], is 200 m, 50 m, and 25 m, respectively, for open, semi-open, and closed office environments. An open environment refers to there being line-of-sight (LoS) connectivity between a transmitter and a receiver. Semi-open and closed environments refer to there being obstructions in the form of modular office partitions (cubicles) and solid walls, respectively. In this nomenclature, our testbed environment would be classified as being open along the hallways where the base stations are located and closed elsewhere. The base stations provide overlapping coverage in portions of the floor, and together cover the entire floor.

3.2 Data Collection

A key step in our research methodology is the data collection phase. We record information about the radio signal as a function of the user's location. As discussed in Section 4, we use the signal information to construct and validate models for signal propagation during off-line analysis as well as to infer the location of a user in real time. We refer to the former as the *off-line phase* and the latter as the *real-time phase*.

Among other information, the WaveLAN NIC makes available the *signal strength* (SS) and *the signal-to-noise ratio* (SNR). SS is reported in units of dBm and SNR is expressed in dB. A signal strength of s Watts is equivalent to $10*\log_{10}(s/0.001)$ dBm. A signal strength of s Watts and a noise power of n Watts yields an SNR of $10*\log_{10}(s/n)$ dB. For example, a signal strength of 1 Watt is equivalent to 30 dBm. Furthermore, if the noise power is 0.1 Watt, the SNR would be 10 dB.

The FreeBSD 3.0 WaveLAN driver extracts the SS and the SNR information from the WaveLAN firmware each time a broadcast packet is received[2]. It then makes the information available to user-level applications via *ioctl* system calls. We used the *wlconfig* utility, which provides a wrapper around the *ioctl* calls, to extract the signal information.

While Windows also obtains the same information from the WaveLAN NIC, we were unable to determine a programmatic way of extracting the information using the NDIS driver interface. This constrained us to have the Windows-based mobile host broadcast packets (*beacons*) periodically and have the base stations (which run FreeBSD) record signal strength information[3]. However, in a production system with many more mobiles than base stations, it may be desirable to have the latter transmit the beacons and the former measure the signal strength.

We wrote a simple application using Tcl/Tk [Ous94] and Perl [Wal96] to control the entire data collection process from the mobile host. The process operates as follows. First, the clocks on the mobile host and the base stations are synchronized (to within the round-trip latency of the wireless link, essentially less than 5 ms). The mobile host then starts broadcasting UDP packets with 6-byte payloads at a default rate of 4 per second. Each base station (*bs*) records the signal strength (*ss*) measurement[4] together with a (synchronized) timestamp *t*, i.e., it records tuples of the form (*t*, *bs*, *ss*). This information is collected both during the off-line phase and the real-time phase.

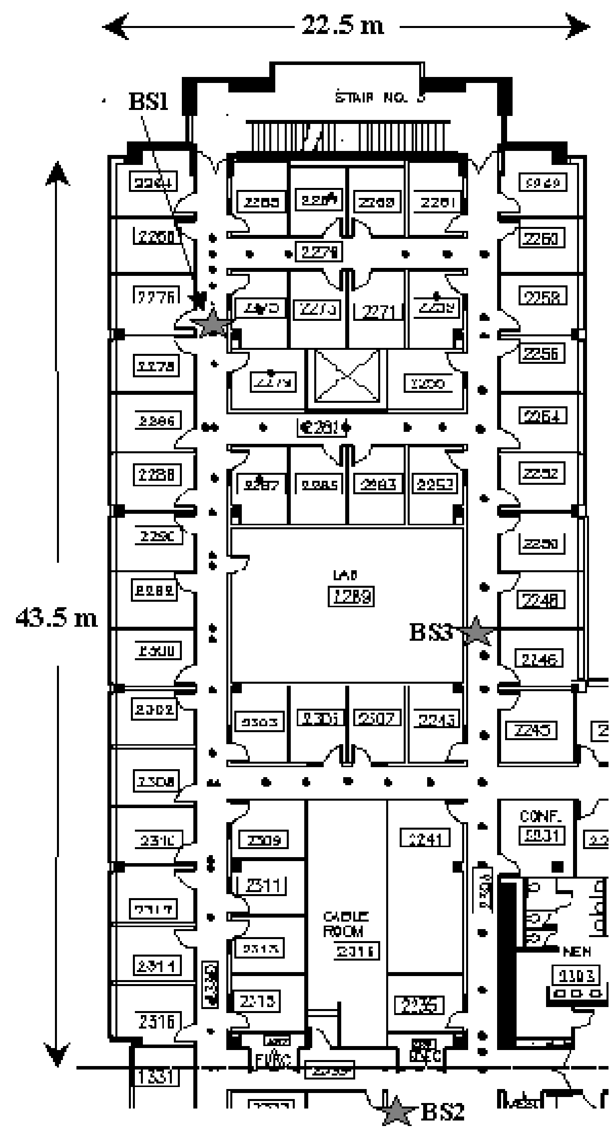

Figure 1 Map of the floor where the experiments were conducted. The black dots denote locations were empirical signal strength information was collected. The large stars show the location of the three base stations. The orientation is North (up) and East (right).

[2] It is quite easy to modify the driver to record information for other packets as well, but we found no reason to do so.

[3] While our analysis does not assume symmetry of signal strength, the few instances where we measured signal strength at both ends indicate little asymmetry.

[4] During the course of our experiments, we discovered that the signal strength is a stronger function of location than the signal-to-noise ratio. The latter is impacted by random fluctuations in the noise process. So we only use signal strength information in our analysis.

During the off-line phase alone (and not the real-time phase), the user indicates his/her current location by clicking on a map of the floor. The user's coordinates (*x*,*y*) are recorded together with a timestamp.

During our experiments, we discovered that signal strength at a given location varies quite significantly (by up to 5 dBm) depending on the user's orientation, i.e., the direction he/she is facing. For instance, in one orientation, the mobile host's antenna may have line-of-sight connectivity to a base station's antenna while in the opposite orientation, the user's body may form an obstruction. So, in addition to user location (*x*,*y*), we also recorded the direction (*d*) (one of north, south, east, or west) that he/she is facing at the time the measurement is made[5]. Thus, the mobile host records tuples of the form (*t*,*x*,*y*,*d*). We discuss the implications of the user's orientation in more detail in Section 4.

In all, during the off-line phase, we collected signal strength information in each of the 4 directions at 70 distinct physical locations on our floor. For each combination of location and orientation (i.e., (*x*,*y*,*d*) tuple), we collected at least 20 signal strength samples.

3.3 Data Processing

We present an overview of the data processing that we performed as a precursor to the analyses discussed in Section 4.

3.3.1 Signal Strength Information

Using the synchronized timestamps, we merged all of the traces collected during the off-line phase into a single, unified table containing the tuples of the form (x,y,d,ss_i,snr_i), where $i \in \{1,2,3\}$ corresponding to the three base stations. For each *(x,y,d)* tuple, we computed the mean, the standard deviation, and the median of the corresponding signal strength values for each of the base stations. For much of our analysis, we use this processed data set (primarily the mean) rather than the original, raw data set.

We wrote routines to search through the processed data set to determine exact as well as closest matches. There is a fair amount of database research literature that describes efficient data structures and algorithms for such multi-dimensional searches (e.g., *R-Tree* [Gut84], *X-Tree* [Ber96], *optimal k-nearest neighbor search* [Sei98], etc.) However, we chose a simple linear-time search algorithm because our relatively small data set and dimensionality (at most 3, as explained in Section 4) did not warrant the complexity of the aforementioned algorithms. Moreover, the focus of our research is on doing the analysis rather than developing an optimal implementation.

3.3.2 Building Floor Layout Information

We obtained the layout information for our floor, which specified the coordinates of each room. We also obtained the coordinates of the three base stations. Using these and the Cohen-Sutherland line-clipping algorithm [Fol90], we computed the number of walls that obstructed the direct line between the base stations and the locations where we had collected the empirical signal strength data. We use this to build an accurate signal propagation model (Section 4.2).

4 Algorithms and Experimental Analysis

In this section, we discuss several algorithms for user location and tracking and present an analysis of how well these perform using our experimental data.

A basic premise of our work is that signal strength information provides a means of inferring user location. To demonstrate that this is a reasonable premise, we show in Figure 2 how the signal strength measured at each base station varies as the user walks along the outer hallway of the floor in a counter-clockwise direction. The walk begins and terminates at the north-west corner (close to BS_1).

From Figure 2, we observe that there is a definite trend in the signal strength measured at the three base stations as the user walks around the loop. Not surprisingly, the signal received at a base station is the strongest when the user is close to it and weakest when the user is far away. This strong trend is an indication that using signal strength to estimate location may be a promising approach.

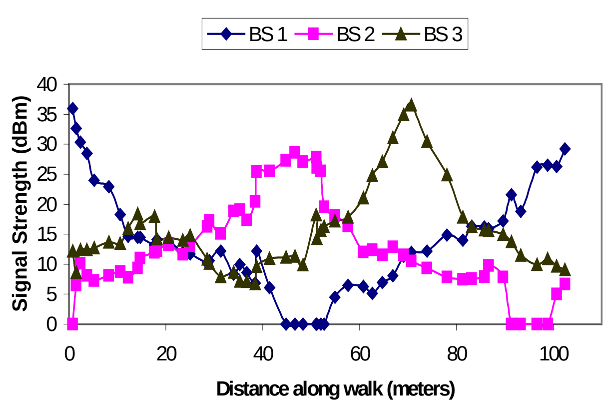

Figure 2 Signal strength recorded at the three base stations as the user walks around the floor.

[5] While there are other sources of fluctuation, such as the movement of other people and objects, these tend to be random. In contrast, the body of the person carrying the mobile host introduces a systematic source of error.

Our basic approach is *triangulation*[6]. Given a set of signal strength measurements at each of the base stations, we determine the location that best matches the observed signal strength data. We then "guess" that to be the location of the user. There are multiple variations of this basic idea that arise because of several choices for each of the following:

- Ways of summarizing the signal strength samples at the base stations.
- Basis for determining the best match.
- Metric for determining the best match.

We discuss each of these in turn.

First, we summarize multiple signal strength samples from a base station using the sample mean. In the case of a static user whose location and orientation are fixed (the *user location* problem), it is clear which signal strength measurements should be included in the sample set. In the case of a mobile user (the *user tracking* problem), it is less clear what the sample set should be. In the latter case, we define the sample set to be all samples that lie within a sliding time window.

Second, to determine the location and orientation that best match a given (summarized) set of signal strength measurements, we first need to determine what the signal strength at each base station should be for a particular combination of user location and orientation. We use a couple of alternative approaches for this purpose. The first is the *empirical method* where we use the location and signal strength data measured during the off-line phase (Section 3.2). The second approach is *signal propagation modeling*. Our model accounts for both free-space loss and loss due to intervening obstructions.

Third, we need a metric and a search methodology to compare multiple locations and pick the one that best matches the observed signal strength. We term our general technique *nearest neighbor(s) in signal space (NNSS)*. The idea is to compute the *distance* between observed set of signal strength measurements, (ss_1, ss_2, ss_3), and the expected signal strength, (ss'_1, ss'_2, ss'_3), at a fixed set of locations and then pick the point that minimizes the distance. In our analysis, we use the *Euclidean distance* measure, i.e., $sqrt((ss_1-ss'_1)^2+(ss_2-ss'_2)^2+(ss_3-ss'_3)^2)$. It is possible to use other distance metrics, for example, the sum of the absolute differences for each base station ("*Manhattan" distance* [Cor90]) or a metric weighted by the signal strength level at each base station. We experimented briefly with these alternatives, but don't present the results here because of space limitations.

In all of our analysis, we characterize the goodness of our estimate of the user's location using the *error distance*, which is the Euclidean distance between the actual location and the estimated location (in physical space).

[6] It is just coincidental that we have three base station in our testbed.

4.1 Empirical Method

In this case, we use the empirical data discussed in Section 3.2 to construct the search space for the NNSS algorithm. We present results of various analyses we performed on this method. Unless otherwise mentioned, we assume the user to be stationary.

4.1.1 Basic Analysis

For the basic analysis, we use all the (more than 20) signal strength samples collected for each of the 70*4 = 280 combinations of user location and orientation. In the analysis, we pick one of the locations and orientations at random, and then conduct an NNSS search for the corresponding signal strength tuple in the space of the remaining 69 points times 4 orientations. This emulates the process of tracking a (stationary) user during the real-time phase.

We compare the empirical method with two other methods: *random selection* and *strongest base station selection* [Hod97]. With random selection, we estimate the user's location by picking one of the 70 points at random, regardless of the signal strength information. With strongest base station selection, we guess the user's location to be the same as the location of the base station which records the strongest signal. A comparison with these simple methods enables us to evaluate how worthwhile the effort expended in our more sophisticated techniques is.

Figure 3 shows the cumulative distribution function (CDF) of the error distance for the empirical, strongest base station, and random methods. The empirical method performs significantly better than both of the other methods. Table 1 summarizes the information in the figure in terms of the 25th, 50th (median), and 75th percentile values of the error distance for each method.

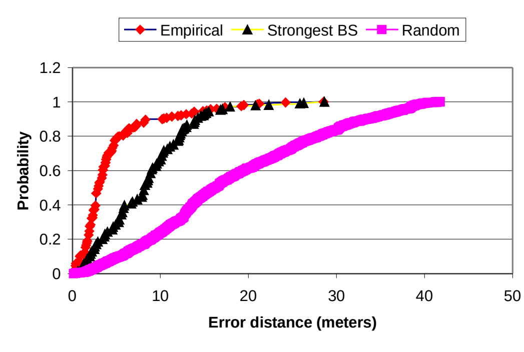

Figure 3 CDF of the error in location estimation (in meters) for three different algorithms.

Considering the median (50th percentile), for instance, the empirical method has a resolution of under 3 meters, which is about the size of an office room in our building. In terms of linear resolution, it is 2.8 times better than the strongest base station method and 5.5 times better than the random method. In terms of spatial resolution, the improvement is even greater: 7.7 and 30.6 times, respectively. We use the percentile values for the empirical method in Table 1 as a basis for comparison in the rest of the analysis.

Method	25^{th} (meter)	50^{th} (meter)	75^{th} (meter)
Empirical	1.92	2.94	4.69
Strongest	4.54 (2.4x)	8.16 (2.8x)	11.5 (2.5x)
Random	10.37 (5.4x)	16.26 (5.5x)	25.63 (5.5x)

Table 1 The 25th, 50th, and 75th percentile values of the error distance. The numbers in parenthesis indicate how much worse the strongest BS and random methods are compared to the empirical method.

In summary, the empirical method performs extremely well. Next, we discuss ways of making it perform even better.

4.1.2 Multiple Nearest Neighbors

Unlike the basic analysis where we only considered the single nearest neighbor in signal space, we now consider *k* nearest neighbors, for various values of *k*. The intuition is that often there are multiple neighbors that are at roughly the same distance from the point of interest (in signal space). Given the inherent variability in the measured signal strength at a point, there is no fundamental reason to pick only the closest neighbor (in signal space) and reject others that are almost as close.

A second, and equally important, reason to consider more neighbors than the single nearest neighbor is that it is likely that the *error vector* (in physical space) corresponding to each neighbor is oriented in a different direction. So averaging the coordinates of the neighbors may yield an estimate that is closer to the user's true location than any individual neighbor is. Figure 4 illustrates this for *k=3* nearest neighbors.

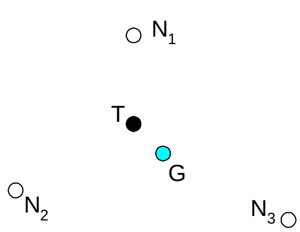

Figure 4 An illustration of how averaging multiple nearest neighbors (N_1, N_2, N_3) can lead to a guess (G) that is closer to the user's true location (T) than any of the neighbors is individually.

Our analysis of the empirical method with averaging over *k* nearest neighbors shows that for small *k*, averaging has some benefit though not very significant. For instance, for *k=5*, the 25^{th} percentile of error distance is 1.5 meters (22% better than the 1.92 meters in Table 1) and the 50^{th} percentile is 2.75 meters (9% better). For large *k*, accuracy degrades rapidly because points that are physically far from the true location also are included in the averaging procedure, thereby corrupting the estimate.

The reason why the benefits of averaging are not very significant even for small *k* is that often the *k* nearest neighbors in signal space are *not* *k* physically distinct points. In many instances, multiple nearest neighbors in signal space correspond to different orientations at the same point in physical space. So averaging in physical space does not improve the location estimate by very much.

4.1.3 Maximum signal strength across orientations

Since the dependence of signal strength on orientation creates challenges for location estimation, we analyze how well the empirical method would perform if orientation were not an issue. For each user location in the trace collected in the off-line phase, we compute the maximum signal strength at each base station across the four possible orientations at that location[7]. Note that the maximum for each base station may correspond to a different orientation. The goal is to emulate the situation where the signal generated by the mobile host is *not* obstructed by the user's body. While this may not be realistic given the antenna design and positioning for existing wireless LANs, it may be possible to approximate this "ideal case" with new antenna designs (wearable antennae akin to wearable computers).

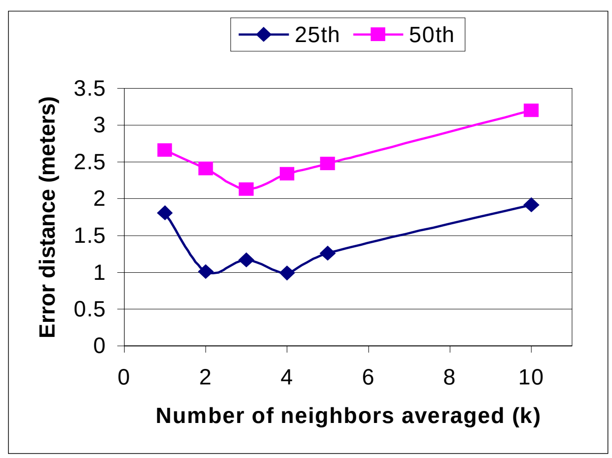

Figure 5 The error distance when the empirical method with averaging is used on the data set containing the maximum signal strength measurement for each location.

[7] Note that for each base station, we first compute the mean over samples for each of the four orientations at a location and then pick the maximum among the four means.

We repeat the analysis of the previous sections with the smaller "maximum signal strength" data set of 70 data points (instead of 70*4=280 data points in the original data set). In Figure 5, we plot the 25^{th} and the 50^{th} percentile values of the error distance with averaging over neighbor sets of various sizes.

We make a couple of observations. First, just as expected, the use of the maximum signal strength data set improves the accuracy of location estimation slightly `even when no averaging is done (*k=1*). The 25^{th} percentile value of the error distance is 1.8 meters and the 50^{th} percentile 2.67 meters, 6% and 9% better, respectively, compared to Table 1. Second, averaging over 2-4 nearest neighbors improves accuracy quite significantly; the 25^{th} percentile is about 1 meter (48% better than in Table 1) whereas the 50^{th} percentile is 2.13 meters (28% better). The reason averaging is more effective in this case is that unlike in Section 4.1.2, the sets of k nearest neighbors in signal space necessarily correspond to *k* physically distinct locations.

4.1.4 Impact of the number of empirical data points

Thus far in our analysis, we have considered signal strength information collected at all 70 physically distinct locations. We now investigate how the accuracy of location estimation is impacted if we only had data from fewer physical locations.

For each value of *n*, the number of physical locations (ranging between 2 and 70), we conducted 20 runs of our analysis program. In each run, we picked *n* points at random from the entire data set collected during the off-line phase and used this subset to construct the search space for the NNSS algorithm. We collated the error distance data from all the runs corresponding to the same value of *n*.

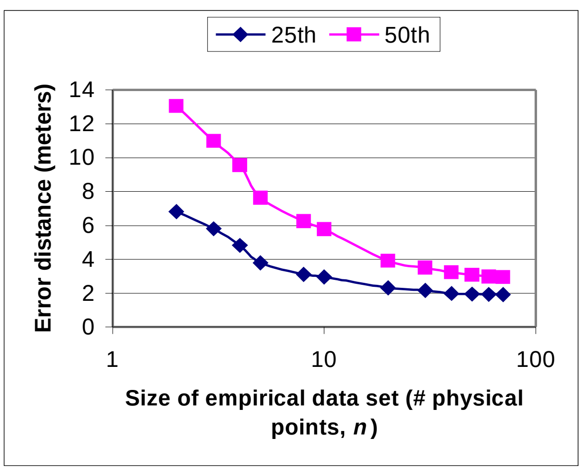

Figure 6 The error distance versus the size of the empirical data set (on a log scale).

For small *n* (5 or less), the error distance is a factor of 2 to 4 worse when the entire empirical set containing 70 physical points is used. But the error distance diminishes rapidly as *n* increases. For *n=20*, the median error distance is within 33% of optimal and for n=40, it is within 10% of optimal. The diminishing returns as *n* gets large is due to the inherent uncertainty in the reliability of the measured signal strength caused by fluctuations. This translates into inaccuracy in the estimation of physical location. So there is little benefit in obtaining an empirical data at a very fine (physical) granularity.

In summary, for our floor, the empirical method would perform almost as well with an empirical data set of 40 physical points as with a set of 70 points. In practice, we may be able to make do with even fewer than 40 points by picking physical locations that are distributed uniformly over the area of the floor rather than at random.

4.1.5 Impact of number of samples at each location

In the analysis presented so far, we have worked with the mean of all of the (20 or more) samples recorded for each combination of location and orientation during the off-line phase. While it may be reasonable to construct the empirical data set with a large number of samples (since it is a one-time task), there may be constraints on the number of samples that can be obtained in real-time to determine a user's location. So we investigate the impact of a limited number of real-time samples (while retaining the entire empirical data set for the NNSS search) on the accuracy of location estimation. Our analysis shows that only a small number of real-time samples are needed to approach the accuracy obtained using all of the samples (Table 1). With just 1 real-time sample, the median of error distance is about 30% worse than when all samples were considered. With 2 samples, it is about 11% worse and with 3 samples it is under 4% worse.

4.1.6 Impact of user orientation

As we have already discussed, the user's orientation has a significant impact on the signal strength measured at the base stations. In Section 4.1.3, we did a best-case analysis using the maximum signal strength across all four orientations. We now consider, in some sense, the worst case where the empirical data set only has data points corresponding to a particular orientation (say north) while the real-time samples correspond to the opposite orientation (i.e., south). We compute the error distance for all four combinations of opposing directions: north-south, south-north, east-west, and west-east.

We observe a fairly significant degradation in the accuracy of location estimation. For instance, for the north-south case, the 25^{th} percentile of the error distance was 2.95 meters (54% worse than in Table 1) while the 50^{th} percentile (median) was 4.90 meters (67% worse). This degradation underscores the importance of obtaining empirical data for multiple orientations to construct the NNSS search space. However, even in this worst case, the empirical method outperforms the strongest base station and random methods by a factor of 2 to 4.

4.1.7 Tracking a mobile user

In this sub-section, we analyze the problem of *tracking* a *mobile* user rather than *locating* a *stationary* user, as we have done so far. For this analysis, we collected a new set of data corresponding to random walks by the user along the hallways of our floor. We collected 4 signal strength samples per second at each of the base stations. Assuming that the user walked at a uniform pace, we determine the true location of the user at each time instant.

We reduce the problem of tracking the mobile user to a sequence of location determination problems for a (nearly stationary) user. We use a sliding window of 10 samples to compute the mean signal strength on a continuous basis. This information is then used with the empirical method to constantly estimate the user's location.

The error distance for tracking the mobile user is only slightly worse than that for locating a mobile user. The median error distance is 3.5 meters, about 19% worse than that for a stationary user.

4.1.8 Summary of Empirical Method

The empirical method is able to estimate user location with a high degree of accuracy. The median error distance is only in the range of 2 to 3 meters, which is of the order of the size of a typical office room. For our experimental environment, much of the accuracy is achieved with an empirical data set containing about 40 physical points and about 3 real-time signal strength samples (at each base station). It is important, however, to obtain data corresponding to multiple user orientations.

The main limitation of the empirical method is that significant effort is needed to construct the signal strength data set for each physical environment of interest (each floor, each building, etc.). Furthermore, the data collection process may need to be repeated in certain circumstances, for instance, when a base station is relocated.

We now discuss a different technique, based on signal propagation modeling, that avoids this limitation.

4.2 Radio Propagation Model

Radio propagation modeling provides an alternative to the empirical method for constructing the search space for the NNSS algorithm.

4.2.1 Motivation

The primary motivation behind the radio propagation model is to reduce the dependence of the user location and tracking algorithm on empirical data. The idea is to determine a mathematical model that characterizes the indoor radio channel. This model is then used to generate a data set of theoretically-computed signal strengths akin to the empirical data set collected in section 3.2. The signal strength information is computed at uniformly spaced pre-determined locations on the floor. The NNSS algorithm can then estimate the location of the mobile user by matching the signal strength measured in real-time to the theoretically calculated signal strengths at these locations. It is clear that the performance this system is directly impacted by the "goodness" of the propagation model and special care is needed in developing this model. In this section we develop the model and the results of doing location determination based on the model.

4.2.2 Determination of the model

For a radio channel, signal propagation in indoor environment is dominated by reflections, diffraction, and scattering of radio waves caused by structures inside the building. The transmitted signal generally reaches the receiver by more than one path, and results in a phenomenon known as *multipath*. Multipath causes fluctuations in the received signal envelope and phase, and the signal components arriving from indirect and direct paths combine to produce a distorted version of the transmitted signal. Since multipath within buildings is strongly influenced by the layout of the building, the construction material used, the number and type of objects in the building, characterizing the radio channel in such an environment is challenging.

We considered three different models and settled on one. The first model we looked at was the well-accepted *Rayleigh fading model* [Has93], which describes small-scale rapid amplitude fluctuation in the absence of a strong received component. The Rayleigh distribution is widely used to describe multipath fading because of its elegant theoretical explanation and the occasional empirical justification. However, in deriving this distribution, a critical assumption made is that all signals reaching the receivers have equal strengths. In general, this is an unrealistic assumption. Our empirical data shows that for a number of sample points (along the hallways), there exists a dominant line-of-sight component that is not accounted for by this distribution. For this reason, we did not use this distribution.

A second model we considered was *the Rician distribution model* [Ric44]. The Rician distribution occurs when a strong path exists in addition to the low level scattered path. This strong component may be the LoS path or a path that goes through much less attenuation than the other arriving components. The Rician distribution contains the Rayleigh distribution as a special case. When the strong path is eliminated, the amplitude distribution becomes Rayleigh. While the model is intuitively appealing, it is very difficult to determine the model parameters (i.e., the local mean of the scattered power and the power of the dominant component) precisely as this requires physically isolating the direct wave from the scattered components. To keep the system simple and easy to deploy, we opted against using this distribution to model the radio channel.

We found a good compromise between simplicity and accuracy in the *Floor Attenuation Factor* propagation model (FAF) suggested by [Sei92]. We like this model because it provides flexibility in accommodating different building layouts while taking into account large-scale path loss. Our measurements confirm earlier findings that signal power decreases exponentially with distance if the attenuation due to the intervening obstacles is compensated for. We adapted the original model proposed by Seidel and Rappaport, which included an attenuation factor for building floors, to disregard the effects of the floors and instead consider the effects of obstacles (walls) between the transmitter and the receiver. The *Wall Attenuation Factor* (WAF) model is described by

$$P(d)[dBm] = P(d_o)[dBm] + 10n\log\left(\frac{d}{d_o}\right) + \begin{cases} nW(WAF) & nW < C \\ C(WAF) & nW \geq C \end{cases}$$

where n is the path loss component that indicates the rate at which the path loss increases with distance. $P(d_o)$ is the signal power at some reference distance d_o and d is the transmitter-receiver (T-R) separation distance. C is the maximum number of obstructions (walls) up to which the attenuation factor makes a difference, nW is the number of obstructions (walls) between the transmitter and the receiver and WAF is the wall attenuation factor. In general the value of n and WAF depends on the building layout, and construction material and is derived empirically. The value of $P(do)$ can either be derived empirically or obtained from the wireless network hardware vendor.

Figure 7 illustrates how the signal strength varies with distance between the transmitter and the receiver. The wide difference in signal strengths between points at similar distances is explained as follows: the layout of the rooms in the building, the placement of base stations, and the location of the mobile user all have an effect on the received signal. Two locations that are at the same distance from the transmitter are affected by different amounts of signal attenuation due to the differences in the number and types of obstructions between them and the transmitter. For example, looking at Figure 7, we notice that two measurements taken at a distance of approximately 36 meters from the transmitter had signal strengths that were 10 dBm apart. One of these two measurements was made by the receiver at a location that had up to 6 walls between it and the transmitter, while the other location had line-of-sight to the base station. Thus it is reasonable to conclude that the number of intervening obstructions effects signal loss at any location and it is possible to classify indoor channels as either LoS or obstructed, with varying degree of clutter.

Previous work in indoor radio propagation modeling has included extensive characterization of signal loss for various materials and at different frequencies [Rap96]. However, using this information in a practical setting is difficult and not as useful since the obstructing materials vary considerably in their physical and electrical characteristics. For example, water causes signal attenuation and the human body is made up of water, so the size of a human body and its orientation can result in different amounts of signal loss. There is no way to characterize such loss precisely since the number and sizes of humans in the building at any particular time is generally a finite but random number.

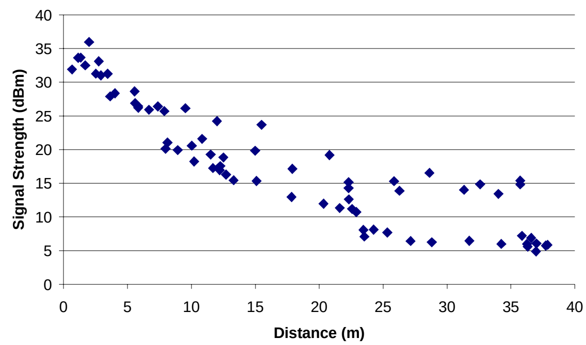

Figure 7 Signal strength as a function of T-R separation derived from the empirical data collected in Section 3.2.

The building in which our research was carried out has a large number of rooms filled with furniture, electronic equipment and people (see Figure 1). In order to account for these obstructions while keeping the model simple, we took the following approach: we combined the effect of signal attenuation due to various obstructions into one number which we called the *wall attenuation factor*. We took the map of the building, and for every location where the signal strength was measured, we determined the number of intersecting walls between the location and the three base stations. As explained in the next paragraph, we determined the value of WAF and applied a correction to the measured signal to compensate for signal loss due to obstructing structures.

In order to determine a suitable WAF, we carried out the following experiment: we measured the signal strength at the receiver when the receiver and the transmitter had line-of-sight. We then measured the signal strength with varying but known number of walls between the receiver and the transmitter. We computed the average of the difference between these values and determined *WAF*. We observed that the amount of attenuation dropped-off as the number of walls separating the transmitter and the receiver increased. This observation is consistent with [Sei92] where the attenuation between different floors was considered and shown to drop-off as the number of floors between the transmitter and the receiver increased. In general, with larger T-R separation and large number of intervening walls, free-space path loss dominates over loss due to obstructions. We chose *WAF* to be 3.1 dBm and C to be 4 (where C represents number of walls that are factored into the model). Figure 8 shows the result after the measured signal strength has been compensated for signal loss due to the intervening walls between the transmitter and the receiver. We observe that the resulting plot shows a trend similar to the free-space loss trend, thus corroborating our intuition that the attenuation factor propagation model is a good way to proceed.

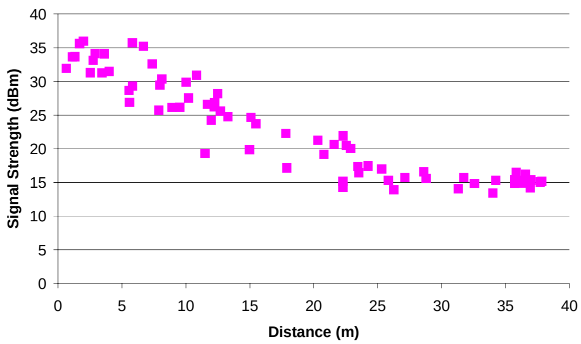

Figure 8 Effect of applying correction for intervening walls between the base station and the mobile user

Once we had taken the effect of walls into account and had created the "corrected" data for all three base stations, we proceeded to determine the other two parameters, (n and P_{do}) of our model. Since the propagation model can be trivially reduced to a form where it exhibits a linear relationship between the theoretical signal strength and logarithm of the distance between the transmitter and the receiver, we applied simple linear regression to determine the parameters of the model [Jai91]

Table 2 contains the numerical values of the model parameters for the three base stations considered separately and when taken together. We note that the values for the path loss exponent (n) and the reference signal strength (P_{do}) for all three base stations are similar despite their different physical locations and surroundings. This result is encouraging since it indicates that the parameter values are not tied to the specific location of the base stations. The slightly higher values of P_{do} are explained on the basis of multipath propagation, which is not taken into account by the model. We observe that the values of the path loss exponent are smaller than what has been reported in previous studies of indoor radio propagation modeling [Rap96]. However, they are consistent with our expectations since we compensate the measured signal strength for attenuation due to obstructions, and since we do not consider multipath (which can boost the signal strength at a given location). R^2 represents the *coefficient of determination*, which is a useful measure for indicating the goodness of regression. The high values of R^2 suggest that there is a strong correlation between the estimated and measured values of the signal strength. Another value of interest shown in the table is the mean squared error. The numbers suggest that we can be confident that the attenuation factor propagation model fits the measured data nicely.

The final column in Table 2 shows the values for P_{do} and n when the data from all the transmitter-receiver pairs was combined. The motivation behind this was to determine a value of P_{do} and n that could be used for all base stations without overly effecting the result. The advantage of using a common value is that it avoids the need for individual measurements of each base station as they are installed in the network, thus greatly reducing the cost of system setup. We can then use these values to estimate signal strength values at various points within the building.

	BS_1	**BS_2**	**BS_3**	**All**
P_{do}	57.58	56.95	64.94	58.48
n	1.53	1.45	1.76	1.523
R^2	0.81	0.65	0.69	0.72
MSE	10.49	13.98	7.34	9.82

Table 2 Model parameter estimates based on linear regression applied to measured data

Figure 9 illustrates how the predicted values of the signal strength generated with the propagation model compared against the actual measurements after they were corrected for the intersecting walls. We observe a good match between the two. While this plot is for one of the three base stations, plots for the other two base stations exhibit similar results.

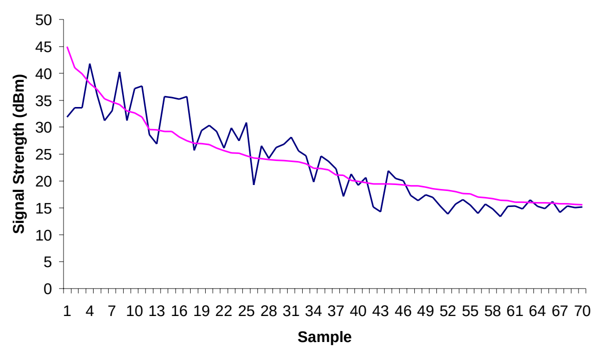

Figure 9 Comparison of predicted versus measured signal strength.

4.2.3 Results of using the Propagation Model

To determine the performance of location estimation with the signal propagation modeling method, we computed the theoretical values of the signal strengths for several locations on the floor. We then provided this data-set to the NNSS algorithm.

Considering the median (50th percentile), the propagation method provides a resolution of about 4.3

meters, compared to a resolution of 2.94 meters for the empirical method and 8.16 meters for picking the strongest base station method (Table 1). For the 25th percentile the propagation method provides a resolution of 1.86 meters compared to 1.92 meters for the empirical method and 4.94 meters for the strongest base station method. While the propagation method is not as good as the empirical method it is significantly better than the strongest BS and random methods. So even without any empirical measurements, our system based on the propagation model alone would perform significantly better than the course level base station method adopted in [Hod97]

4.2.4 Summary of Radio Propagation Method

The radio propagation model is a cost effective way of doing location tracking in an indoor wireless radio network. The model is cost effective in the sense that it does not require detailed empirical measurements for location tracking and consequently has a low set up cost. A significant result from Section 4.2.2 is that the parameters for the wall attenuation propagation model are similar across base stations despite their being in different locations. This suggests that the entire system can be located in a different part of the building and the same parameter values used to model propagation and thereby determine location.

5 Discussion

In this section, we discuss some extensions to basic algorithms presented in this paper. We also discuss some applications that are enabled by the user location and tracking algorithms we have developed.

5.1 Self-Adaptive Kalman Filtering

In a previous paper we have shown that continuous processing of signal strength measurements enables prediction of a mobile user's immediate future location [Anon1]. Our research objective in that paper was to build a system to predict the speed and trajectory of a high-speed mobile user in a wide-area cellular network. The input to that system, instantaneous noisy signal strength measurements from the neighboring base stations, is identical to the input to the system described in this paper. We built a two-level hierarchical location prediction (HLP) algorithm consisting of approximate pattern matching and classical stochastic signal processing techniques. While the pattern-matching component required historical movement patterns to yield good results when the user exhibited regular movement patterns, the signal processing technique required no historical movement data. We showed that a self-adaptive extended Kalman filter could take noisy measurements at the input and still provide an accurate prediction for next location of the mobile.

The HLP algorithm can be employed in an indoor local-area wireless network, and we expect that it would improve the location tracking results. However, the computational complexity of determining the Kalman gain matrix with each new sample is high. It is not clear that the amount of additional accuracy obtained from applying this computationally complex algorithm would be worth while. In particular, the additional accuracy might not enable applications that require precise location determination and for the ones that don't depend on such precise location estimation, the accuracy provided by the algorithms proposed in this paper may well be enough.

5.2 Applications Enabled

Different location-aware services have varying requirements for the accuracy of location information. Our experimental results indicate the feasibility of estimating user location to within a couple of meters. We believe that this capability enables a rich class of location-aware service. For example, our system can be used to locate nearby network resources such as printers, fax machines, copy machines etc. The mobile user could expect to get good result to the request *"print this document on the closest printer"*. The system would determine the location of the user, find a nearby printer with the right attributes, graphically indicate to the user the location of this printer, and then print the document on this printer. Similarly a request like *"Show me the map of the area I am in"* will result in the appropriate map being displayed on the screen. Going in the reverse direction, queries such as *"Where is user A?"* would be answered by the system and the location of the mobile would be returned to within a certain range of his or her actual location.

6 Conclusions

In this paper, we have addressed the problem of user location and tracking in an in-building RF network. Our algorithm is based both on empirical signal strength measurements and a simple yet effective signal propagation model. While the empirical method is superior in terms of accuracy, the signal propagation method generalizes to a much greater extent.

We have shown the despite the hostile nature of the radio channels, we are able to locate and track a user with a high degree of accuracy. Based on experiments using the WaveLAN network, we determined the median resolution of location estimation to be in the range of 2 to 3 meters, which is about the size of a typical office room.

Our results give us confidence that it is possible to build an interesting class of location-aware services, based solely on the RF in-building wireless data network, thus adding value to such a network. This, we believe, is a significant contribution of our research.

References

[ATC98] http://www.ascension-tech.com

[Azu93] R. Azuma, "Tracking Requirements for Augmented Reality," *Communications of the ACM*, Vol. 36, No. 7, pp: 50-51, July 1993

[Ber96] S. Berchtold, D. A. Keim, H. P. Kriegel, "The X-tree: An Index Structure for High-Dimensional Data,", Proc. VLDB, 1996

[Cor90] T. H. Cormen, C. E. Leiserson, R. L. Rivest, "Introduction to Algorithms," The MIT Press, 1990

[Fol90] Foley, van Dam, Feiner, and Hughes, Computer Graphics Principles and Practice, Second Edition, 1990

[For95] S. F. Fortune, D. M. Fay, . W. Kernighan, O. Laudron, R. A. Valenzuela, and M. H. Wright, "WISE Design of Indoor Wireless Systems," IEEE Computational science and Engineering, pp. 58-68, Spring 1995

[GPS99] P. Enge, and P. Misra, "Special Issue on GPS: The Global positioning System," *Proceedings of the IEEE*, pp. 3-172, January 1999

[Gut84] A. Guttman, "R-Trees: A Dynamic Index Structure for Spatial Searching," ACM, 1984

[Har94] A. Harter and A. Hopper, "A Distributed Location System for the Active Office," *IEEE Network*, January 1994

[Has93] H. Hashemi, "The Indoor Radio Propagation Channel," *Proceedings of the IEEE*, Vol. 81, No. 7, pages 943-968 July 1993

[Hod97] T. D. Hodes, R. H. Katz, E. S. Schreiber, and L. Rowe, "Composable Ad Hoc Mobile Services for Universal Interaction," *MobiCom '97 Proceedings*, pp: 1-12, September 1997

[Jai91] R. Jain, The Art of Computer Systems Performance Analysis, *John Wiley*, 1991

[Maa97] H. Maass, "Location-Aware Mobile Applications based on Directory Services," *MobiCom '97 Proceedings*, pp. 23-33, September 1997

[Nel98] G. J. Nelson, "Context-Aware and Location Systems," Ph.D. Theses, University of Cambridge, Computer Laboratory, Cambridge, UK., January 1998

[Ous94] J. K. Ousterhout, "Tcl and the Tk Toolkit," Addison-Wesley Professional Computing Series, 1994

[Ric44] S. O. Rice, "Mathematical analysis of Random Noise," *Bell Systems Technical Journal,* Vol. 23, pp. 282-332, 1944, and vol. 24, pp. 46-156, 1954

[Roa96] "Digital RoamAbout 915/2400 DS/PC Card and ISA Network Adapter: Installation and Configuration," Digital Equipment Corporation, April 1996

[Rap96] T. S. Rapport, Wireless Communications – Principles and Practice, *IEEE Press*, 1996

[Sei98] T. Seidl, H. P. Kriegel, "Optimal Multi-Step k-Nearest Neighbor Search," Proc. ACM SIGMOD, 1998

[Sei92] S. Y. Seidel and T. S. Rapport, "914 MHz path loss prediction Model for Indoor Wireless Communications in multi-floored buildings," *IEEE Transactions on Antennas and Propagation*, Vol. 40, No. 2, pp. 207-217, February 1992

[Tek98] S. Tekinay, "Wireless Geolocation Systems and Services," Special Issue of the *IEEE Communications Magazine*, April 1998

[Wal96] L. Wall, T. Christiansen, R. L. Schwartz, "Programming Perl," O'Reilly & Associates, Inc., 1996

[Wan92] R. Want, A. Hooper, V. Falcao, J. Gibbons., "The Active Badge Location System," *ACM Transactions on Information Systems*, Vol. 40, No. 1, pp. 91-102, January 1992

[Anon1] Reference withheld to preserve anonymity