ON INDOOR POSITION LOCATION WITH WIRELESS LANS

P. Prasithsangaree[1], P. Krishnamurthy[1], P.K. Chrysanthis[2]

[1] Telecommunications Program, University of Pittsburgh, Pittsburgh PA 15260, {phongsak, prashant}@mail.sis.pitt.edu
[2] Department of Computer Science, University of Pittsburgh, Pittsburgh PA 15260, panos@cs.pitt.edu

Abstract – Location aware services are becoming attractive with the deployment of next generation wireless networks and broadband multimedia wireless networks especially in indoor and campus areas. To provide location aware services, obtaining the position of a user accurately is important. While it is possible to deploy additional infrastructure for this purpose, using existing communications infrastructure is preferred for cost reasons. Because of technical restrictions, location fingerprinting schemes are the most promising. In this paper we are presenting a systematic study of the performance/tradeoff and deployment issues. In this paper we present some experimental results towards such a systematic study and discuss some issues related to the indoor positioning problem.

I. INTRODUCTION

Mobile computing has evolved over the last several years and has aimed at providing mobile users anytime, anywhere access to the right information at the right time. The position location of users is an important component of mobile computing to assist them with their desired goals, and making either the smart workplace or home really meaningful. Knowledge of the positions of users combined with user profiles could significantly help in network planning, in load balancing, caching of information closer to the user, radio resource management and designing other performance enhancement methods. Position location is also receiving increased importance for public safety issues.

It is possible to obtain the position location of a mobile station (MS) in two ways: by using a special infrastructure for positioning such as the global positioning system (GPS) or by enhancing the existing communications infrastructure to determine the location of users. GPS is not suitable for indoor areas because of the lack of coverage and it is very expensive in terms of labor, spectrum and capital costs to implement a specialized infrastructure in indoor areas solely for position location. As such, it is preferable to employ the existing wireless communications infrastructure to determine the location of users within the network. In indoor areas, the wireless communications infrastructure is primarily based on wireless local area networks (WLANs), in particular the IEEE 802.11b standard that supports raw data rates of 11 Mbps [1]. Our focus here thus lies on experimental results with an IEEE 802.11b WLAN.

With existing communications, there are three basic methods for determining the location of users: (a) triangulation that requires at least three distinct estimates of the distance of the MS from known fixed locations[1], (b) using the direction or angle of arrival (AOA) of at least two distinct signals from known locations and (c) employing *location fingerprinting* schemes. Owing to the harsh multipath environment in indoor areas, techniques that use triangulation or direction are not very attractive and often can yield highly erroneous results [2-4]. Location fingerprinting refers to techniques that match the *fingerprint* of some characteristic of the signal that is location dependent. The fingerprints of different locations are stored in a database and matched to measured fingerprints at the current location of an MS. Some companies such as [5] have used the multipath characteristics of a signal as its fingerprint. Such techniques require specialized hardware in every base station (BS) (or access point - AP) to correlate the multipath characteristics. In WLANs, an easily available signal characteristic is the received signal strength (RSS) and this has been used in [8-12] for fingerprinting. The RSS is a highly variable parameter and issues related to positioning systems based on RSS fingerprinting are not understood very well.

There is a significant cost in the comparison of measured data with the stored information. This may not very costly for small areas, but it becomes an increasingly important component as the area to be covered and the number of users becomes large. However, due to the physical and technological limitations associated with other techniques, location fingerprinting schemes remain the most feasible solution for indoor position location. While the work in [6-9] addresses the issue of accuracy of using location fingerprinting, very little has been done to systematically approach this problem. The tradeoffs between deployment issues, accuracy, size of the database, robustness of the system if some access points fail, and performance are lacking. In this paper we discuss these issues and present some preliminary experimental results on the deployment issues, performance, accuracy and robustness of location fingerprinting using the RSS. Our analysis is based on data collected in our building where an IEEE 802.11b WLAN is deployed.

In Section II, we elaborate on position location issues and provide a quick overview of related work. Our experimental testbed, methodology, data collection and the creation of the database are described in Section III. Section IV discusses the analysis and experimental results.

[1] Distance estimates can be obtained from the times / time differences of arrival of signals (using the speed of light) or from the RSS that falls as a function of distance.

II. Issues in RSS Based Location Fingerprinting

A. RSS-based Techniques

The big advantage of RSS-based techniques is that we can use the existing infrastructure to deploy a positioning system with minimum additional devices. It is far easier to obtain RSS information than the multipath characteristic, the time or angle of arrival that require additional signal processing.

The RSS information can be used to determine the distance between a transmitter and a receiver in two ways. The first approach is to map the path loss of the received signal to the distance traveled by the signal from the transmitter to the receiver. With the knowledge of the RSS from at least three transmitters, we can locate the receiver by using triangulation [3]. Here, there is no database search and the positioning delay is just related to the communication and computation. However, inside a building, the variation of the RSS with distance (the inaccuracy of the path-loss model) is significant due to obstructions and multipath fading effects. As such it is usually not reliable to use the RSS in this manner. This method is used in [9] along with training and interpolation to improve accuracy. Another way is to use the path loss models to compile an artificial database of RSS values [6]. The measured RSS values are then compared with this database to obtain the MS's location. This seems counter-intuitive as it retains the inaccuracy of the path-loss models and the delay associated with the database search.

In order to employ RSS-based techniques with greater accuracy, the second approach is to use the RSS in a fingerprinting scheme. In [6], a system called RADAR that consists of three Pentium-based PCs as access points and a laptop computer as an MS is installed in a building. The three access points (APs) measure the RSS from the client. This is thus a *remote positioning* system where a central point computes the location of the MS. The measurements are then correlated with entries in a database that are filled up with similar measurements performed with MSs at known locations. These measurements consist of average values of the RSS (the average computed using several samples at the known location). In summary the RSS values at three APs are used as the *location fingerprint*. Privacy can be a problem as the APs can determine a client's location even when the client does not want anyone to know his/her location. Additionally, when the number of mobile clients increases, the BSs could be overloaded.

The experiments in the RADAR system are set up only in the hallway where the RSS is stable and strong. The measurements are all in a single floor with three APs. A problem with the multifloor environment is that two or more very different locations could potentially have the same RSS location-fingerprint. In [8, 10], instead of storing *average* RSS values, the joint or marginal distributions of the RSS are used in fingerprinting locations. The experiments are once again conducted in hallways and the accuracy ranges between 5 and 20 feet in most cases.

B. Issues in RSS-based Location Fingerprinting

There are several issues in location fingerprinting based on the RSS that need further evaluation [11]. These are briefly discussed below. We elaborate on many of these issues in our results and expect to continue work in those that are not adequately addressed in this work.

1. Self-positioning: Instead of three AP's measuring the RSS from a MS, the MS could measure the RSS values from multiple APs. As and when it requires its location, it can request the information from the network. This improves the privacy of the MS and reduces the continuous computational burden on the network. We implement a self-positioning mechanism in our work.

2. Granularity: The database entries are collected on a grid of points within the building. The spacing between grid points influences the granularity of the position estimate. Decreasing the spacing (e.g. taking RSS measurements every foot) will increase the database size but are unlikely to yield a better accuracy because the RSS values measured a foot apart will be more or less the same. On the other hand, if the spacing is very large, it may reduce the search space but drastically decrease the accuracy. We consider two specific grid spacings: 5 ft (1.5 m) and 10 ft (3 m) to evaluate the tradeoffs between performance and delay.

3. Algorithms: There are two basic algorithms for computing the location of the MS. These are both based on the *signal distance* between the measured fingerprint and fingerprint entries in the database. The *generalized weighted L_p distance* between a measured RSS vector $[x_1\ x_2\ ...x_N]$ and a database entry $[x_1'\ x_2'\ ...x_N']$ is given by:

$$L_p = \frac{1}{N}\left(\sum_{i=1}^{N}\frac{1}{w_i}\left|x_i - x_i'\right|^p\right)^{1/p} \tag{1}$$

The two algorithms are:

a) Choose the location corresponding to the fingerprint with the minimum distance to the measured fingerprint. The Manhattan L_1 distance ($p = 1$) and the Euclidean L_2 distance ($p = 2$) are considered with $w_i = 1$ for all entries and measurements in related works like [6].

b) Choose M closest database entries (those with the smallest signal distance) and estimate the location based on the average of the coordinates of these M points.

In addition to these algorithms, intuitively, we may consider enhancing the algorithms to take into account the way the data was collected or using the information resulting from the search. The weight w_i can be used to bias the distance by a factor that could indicate how reliable a database entry or a RSS measurement is. Since the database entries are average values of the RSS at a location, the number of samples used to compute this average could be an estimate of the reliability of the database entry. We evaluate such modifications to algorithms and our results surprisingly indicate that the accuracy improvement is negligible.

4. Fault-tolerance: An important issue with position location mechanisms will be their reliability. In many cases, some access points may be disabled because of local power failures, management, upgrades etc. In such cases, the user must still be able to obtain some position location service. In this paper we shut down the AP closest to a MS and evaluate the algorithms used for positioning to determine their robustness in terms of accuracy with an AP unavailable.

5. Performance: Issues related to the performance [12] of positioning are very important. The *accuracy* of the location information (the error in the estimated location of the MS), the *delay* in making the position location estimation (delay in making the position location and conveying this information to the requested party), the *capacity* (how many requests for location estimation can be processed in unit time) and *coverage* (the area where the position location service is available) are important performance measures. The architecture of the building could result in layers of accuracy - for instance, accurate to 5′ in certain areas, accurate to 10′ outside of these areas but within other limits and accurate to 20′ outside these limits. We address these issues in this paper.

III. Experimental Testbed

The experimental testbed is located in our eight-story building. The WLAN consists of ten access points located opportunistically on multiple floors. Consequently, this is a multi-floor experiment although in this work we conducted our experiments only on locations in the 4th floor of the building. The layout of the 4th floor is shown in Figure 1. The floor has dimensions of 45 feet by 105 feet and includes more than 10 rooms. The APs are shown as satellite dishes in this floor plan. The other APs that can be seen from these locations are on the third, fifth and sixth floors.

The data collection to populate the database for correlating the location fingerprint consisted of recording the RSS from each AP as a function of the location. The number of APs seen by the mobile host can vary depending on the RSS, path loss, interference, and multi-path fading at different locations, and this creates the location fingerprint. We measure the RSS at various locations such as in a hallway and in various rooms. These locations have very different fingerprints. In the hallways (where APs are located), the RSS is strong. Therefore, the MS can receive signals from many different APs (which can be considered as a good fingerprint). On the contrary, the RSS is weaker in the rooms, and the MS may not have a good fingerprint. At each location, we calculated the average of 40 samples. The 40 samples are composed of 5 samples in each of the MSs orientation at two different times of the day. We use four orientations at each location as explained in [6]. This is because the RSS at a given location can vary up to 5 dB due to the MSs orientation. We used two very different times of day when the presence or absence of people in the building significantly affects the RSS values. In summary, our database contains fingerprints for 60 locations. We created two databases with different resolutions of 10′ and 5′ feet, and their total sizes are about 20 KB and 50 KB respectively. The MS's location is then calculated from the RSS information received from all available APs that are seen by it. The calculation process involves matching the RSS information with that in the database. For matching, we used a simple a linear search algorithm since our database size is rather small. However, for a larger database, we need search mechanisms that utilize efficient access methods. While this is an important issue, we leave the database performance to future work.

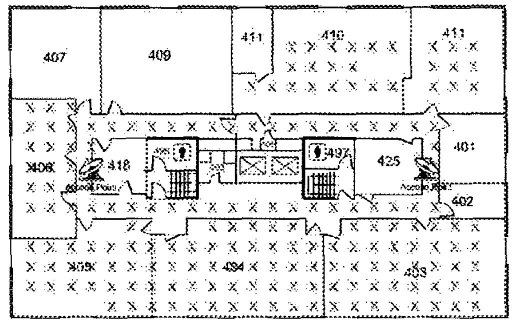

Figure 1: Plan of the 4th Floor and grid of locations

We first match the AP IDs of the RSS information with the AP IDs in the database. If a match is found, the *signal distance* is calculated with the two algorithms as described in Section II. We used the weighted L_p distance where $[x_1\ x_2\ ...x_N]$ is the real-time RSS measured by the MS (x_i is the RSS from AP_i) and $[x_1'\ x_2'\ ...x_N']$ is the RSS vector in the database corresponding to those very same AP. Measurements were not taken in a few rooms in Figure 1. This is not because the RSS values were unavailable, but because access to these rooms is restricted.

IV. Analysis and Experimental Results

A. Algorithms and accuracy

The RSS average tends to vary a lot and it is tempting to use some enhancements to the basic distance measure. We have avoided the probabilistic method as this tends to increase the database size. Increasing the value of p in the L_p distance fails to increase the accuracy. So, we used two weighting schemes, namely the number of signal samples (NSS) and the standard deviation (SD) of the RSS samples used to compute the average RSS in the database as a measure of how reliable the database entry was. We used the NSS and SD values in (1) for w_i. The idea here was to see whether simple techniques like these can improve the accuracy of the estimate of the position without resorting to complicated distribution matching schemes. For these algorithms (referred to as NSS-weight and SD-weight in Figure 2), we used $p = 2$. These algorithms belong to the first category – the search returns the location in the database that has the smallest weighted signal distance to the measured RSS vector.

In contrast, the second type of algorithm (see section II) returns M estimated locations after searching database corresponding to the M smallest Euclidean distances ($p = 2$ and $w_i = 1$ for all i). From the M locations, the location is estimated as the "average". This average corresponds to the centroid of the M locations. For example, if we return the two closest locations ($M = 2$ and called 2-best in Figure 2), the average location is the midpoint of the straight line between the two locations. If three closest locations are returned ($M = 3$, 3-best), the average is the centroid of the triangle formed by the three locations.

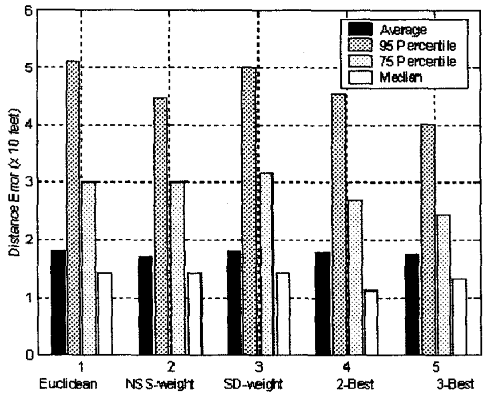

Figure 2: Performance of different algorithms

Figure 2 shows the accuracy of these algorithms. It is quite clear that there is no improvement in weighting the distance with the SD or NSS values. Using the centroid has a better 95% performance compared to all other schemes.

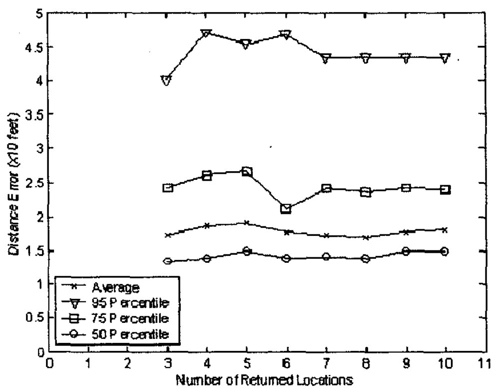

Figure 3: Performance with Closeness-elimination

We also evaluated what we call "the closeness elimination scheme" i.e., we return $K > 3$ locations from our database. We compute the distances between each pair of points and eliminate those that are farthest in terms of these distances, keeping only the three with the closest distances between each other. We then compute the "average" as in the case of the second algorithm and estimate this as the location. In Figure 3, the error in the estimates is shown as a function of the number of returned locations K. Returning more than 8 points tends to induce more error because the 'average' location of points tends to shift from the MSs' correct location due to the closeness elimination. The more the number of returned points, the greater is the tendency for them to be scattered. This scheme performs at the same level as the one with the three closest locations (3-best).

B. Performance-Accuracy Tradeoffs

As discussed in Section II, we also consider the tradeoff between accuracy and performance by increasing the granularity of the grid in the database. Table 1 shows the accuracy (in terms of the mean distance error) and search times associated with databases having a grid spaced 10′ apart and 5′ apart. Increasing the granularity of the grid ***only slightly*** improves the positioning accuracy. However, the size of database entries is increased by about three times, and search time for matching the fingerprint is also significantly increased as expected.

Table 1: Performance Vs Accuracy Tradeoff

Database Granularity	Five Feet	Ten Feet
Average Distance Error (x 10 ft)	2.17	2.5
Time to obtain match (second)	10.43	1.27

C. Fault tolerance of location fingerprinting

Our third experiment was to evaluate the fault-tolerance of the algorithms in the case where one of the APs on the 4th floor is disabled (when the MS is measuring the RSS for position location). This AP incidentally provides the best RSS to most locations. The solid lines show the error without failure and the dashed lines with the AP shut down.

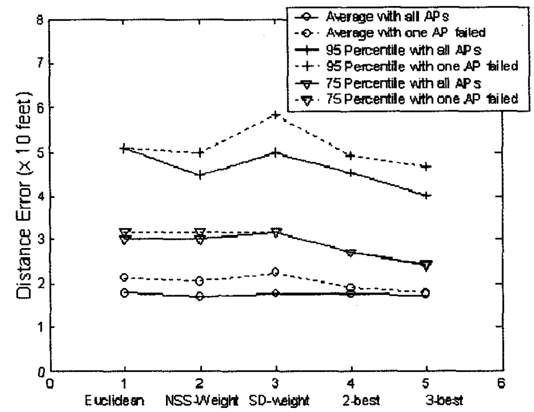

Figure 4: Performance with the failure of an AP

As shown in Figure 4, the average distance error when one AP is down is slightly increased compared to the case with no AP failure. As before, the algorithm that uses the centroid of the 3 closest matches tends to perform better than the other algorithms even with the AP down. This indicates that the position location is quite robust to failures. The reason for the robustness is the fact that there are few other entries in the database that match the location of the MS.

D. Deployment Issues

An important unanswered question with the studies on indoor position location is how the positioning system should be deployed – how many access points are required for a given level of accuracy (what coverage in terms of accuracy does an AP provide), how close the grid spacing should be (having a finer granularity in certain areas could increase accuracy), whether there should be a mixture of algorithms, and whether it is feasible to provide the required accuracy by performing some tradeoffs.

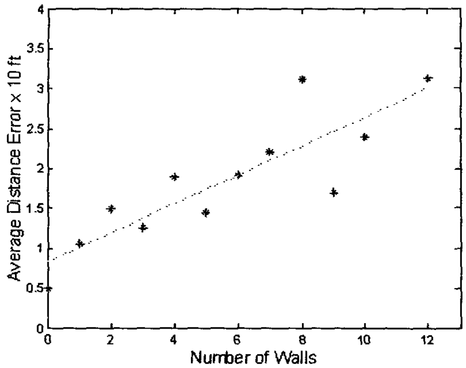

Figure 5: Performance Vs Architecture

In order to gain some insight into answering these questions, we need to correlate the performance measures (accuracy, search time etc.) to the physical architecture of the building. Every location in the building is separated from each AP by some (or no) walls. This number can be obtained by determining how many walls intersect the straight line joining a location and an AP. For each location, we have n_i walls separating it from the i-th AP. Let n_{min} be the minimum over all visible APs at the location (n_{min} = Min$_i$ $\{n_i\}$). We group locations by n_{min} and compute the average error associated for such locations as a function of n_{min}. We use the Euclidean distance in this case. In Figure 5, we show how the average of distance error increases for locations with larger values of n_{min}. The results indicate the problem of user positioning in an indoor area where there is a number of obstructions. As the number of obstructing walls increases, the accuracy of the positioning technique is decreased. The slope of the straight-line fit to the error values is 0.18. This means that for every additional wall, there is approximately 1.8 ft increase in the error in the position estimate. Consequently, to provide a required accuracy in a building, APs may have to be placed in such a way, not simply to provide communications coverage, but also to satisfy the positioning coverage. In our case, two APs can cover one floor of the building for communications, but four APs are required to keep the average distance error to less than 20 feet.

In order to assess the effect of mixed granularity, we used a hybrid database which has a 5′ grid resolution in some areas and 10′ resolution in other areas. Table 2 shows the results of this experiment. It is possible to slightly improve the accuracy of the position error by increasing the granularity of the grid in rooms (where the number of obstructions between the APs and MS locations is larger). As previously observed, there is no improvement by increasing the granularity in hallways where the number of obstructions between APs and MSs is smaller. The conclusion here is that hybrid schemes do provide better accuracy and also good performance by reducing the search time.

Table 2: Using a hybrid database

Database Grid Resolution	10 feet		5 feet	
	Average error x 10′	Search Time (s)	Average error x 10′	Search Time (s)
Only Hallways	1.40	0.32	1.53	1.23
Only Rooms	2.49	0.23	2.16	3.98
Both Areas	2.31	3.77	2.26	11.78
Hybrid 1[1]	2.41	2.84		
Hybrid 2[2]	2.26	6.98		

[1] 5′ grid in hallways and 10′ grid in rooms
[2] 5′ grid in rooms and 10′ grid in hallways

V. Future Work

This paper has provided a framework for systematically analyzing indoor position location using WLANs. We have addressed some of the issues, but more data need to be collected and analyzed to establish models and methodologies for deploying a positioning system. We only have preliminary results on the tradeoffs in the positioning system and how they relate to the physical building architecture. We are investigating further these issues – how many access points are required to provide a given accuracy with a given granularity of the grid in the database, how the database should be organized for better searching speed, whether matching distributions of RSS as in [10] in locations that are severely obstructed from APs is preferable to simply matching the average RSS, etc.

Acknowledgements

The authors would like to thank Sohail Hirani for collecting the extensive RSS data used in this work. We also acknowledge NSF grants IIS-9812532 and EWF-0081327 for partially funding this effort.

References

[1] IEEE P802.11 Working Group Changes and Additions to IEEE 802.11, 1999 Edition. January 2000.

[2] J. Caffery Jr., G. Stuber, "Overview of Radiolocation in CDMA Cellular Systems", *IEEE Comm. Mag*, April 1998.

[3] K. Pahlavan and P. Krishnamurthy, *Principles of Wireless Networks: A Unified Approach*, Prentice Hall PTR, 2002.

[4] K. Pahlavan, P. Krishnamurthy and J. Beneat, "Wideband radio propagation modeling for indoor geolocation applications", *IEEE Comm. Mag*, pp. 60-65, April 1998.

[5] U.S. Wireless Corp. website at http://www.uswcorp.com

[6] P. Bahl and V. N. Padmanabhan, "RADAR: An In-Building RF-based User Location and Tracking System" *Proc. IEEE INFOCOM 2000*, Vol. 2, pp. 775-784, March 2000.

[7] P. Bahl and V. N. Padmanabhan, "Enhancements to the RADAR user location and tracking system", *Technical Report* MSR-TR-2000-12, Microsoft Research, 2000.

[8] M.A. Youssef et al., "A probabilistic clustering-based indoor location determination system", *Technical Report CS-TR-4350 and UMIACS-TR-20002-30*, University of Maryland, 2002.

[9] A. Smailagic et al., "Location sensing and privacy in a context aware computing environment", *Proc. Pervasive Computing*, 2001.

[10] P. Castro et al., "A probabilistic location service for wireless network environments", *Proc. Ubiquitous computing*, Sept. 2001.

[11] P. Krishnamurthy, "Position Location in Mobile Environments", *NSF Workshop on Context-Aware Mobile Database Management (CAMM)*, Providence, RI, Jan. 2002.

[12] S. Tekinay, E. Chao, R. Richton, "Performance benchmarking for wireless location systems", *IEEE Comm. Mag.*, April 1998.