WLAN Location Determination via Clustering and Probability Distributions

Moustafa A. Youssef, Ashok Agrawala, A. Udaya Shankar
Department of Comupter Science and MIND Lab
University of Maryland
College Park, Maryland 20742
{moustafa,agrawala,shankar}@cs.umd.edu

Abstract

We present a WLAN location determination technique, the Joint Clustering technique, that uses (1) signal strength probability distributions to address the noisy wireless channel, and (2) clustering of locations to reduce the computational cost of searching the radio map. The Joint Clustering technique reduces computational cost by more than an order of magnitude, compared to the current state of the art techniques, allowing non-centralized implementation on mobile clients. Results from 802.11-equipped iPAQ implementations show that the new technique gives user location to within 7 feet with over 90% accuracy.

1. Introduction

As ubiquitous computing becomes more popular, the importance of context-aware applications increases. This in turn fuels the need to determine user location, with which the system can provide location-specific information and services [6].

Many systems over the years have tackled the problem of determining and tracking the user position. Examples include GPS [7], wide-area cellular-based systems [16], infrared-based systems [18, 3], ultrasonic-based systems [12], various computer vision systems [9], physical contact systems [11], and radio frequency (RF) based systems [4, 20, 14, 5, 13, 10, 8]. Of these, the class of RF-based systems that use an underlying wireless data network [4, 20, 14, 5, 13, 10], such as 802.11, to estimate user location has gained attention recently, especially for indoor application. Unlike infrared-based systems, which are limited in range, RF-based techniques provide more ubiquitous coverage and do not require additional hardware for user location determination, thereby enhancing the value of the wireless data network.

RF-based systems usually work in two phases: *offline* training phase and *online* location determination phase. During the offline phase, the signal strength received from the access points at selected locations in the area of interest is tabulated, resulting in a so-called *radio map*. During the location determination phase, the signal strength samples received from the access points are used to "search" the radio map to estimate the user location.

RF-based systems need to deal with the noisy characteristics of the wireless channel. Those characteristics cause the samples measured in the online phase to deviate significantly from those stored in the radio map, thereby limiting the accuracy of such systems. Moreover, in order to preserve user privacy and to make the location system scalable, the location determination code should be run on the mobile unit. Since mobile devices are energy-constrained, it is important to reduce the computational requirement of the location determination system.

In this paper, we present an *accurate* and *scalable* system for determining the user location with low computational requirements in an 802.11 wireless LAN (WLAN) framework. The system has two key features: (1) It uses probability distributions to enhance accuracy and tackle the noisy nature of the wireless channels. (2) It uses clustering of map locations to reduce the computational requirements. We call our technique the *Joint Clustering* (JC) technique.

We have evaluated the system in an indoor space spanning a 20,000 square foot. Results obtained show that the Joint Clustering technique gives the user location with over 90% accuracy to within 7 feet with very low computational requirements.

Radio map-based techniques can be categorized into two broad categories: deterministic techniques [4, 14] and distribution-based techniques [20, 5, 13, 10]. Our work lies in the second category. However, none of the previous systems take into account the computational burden of the location determination algorithm. Our work is unique in introducing clustering of radio map locations as an approach to reduce the computational requirements of the location determination techniques and increase the scalability of the system.

The rest of the paper is organized as follows. In Section 2, we describe the noisy characteristics of the wireless channel. Section 3 presents the details of radio map construction and location estimation with the Joint Clustering technique. In Section 4, we describe the evaluation of the techniques in the indoor space and the obtained results. Finally, Section 5 concludes the paper and gives directions for future work.

2. Wireless Channel Characteristics

In this section, we describe our sampling process and the noisy characteristics of the wireless channel which makes the problem of WLAN location determination a challenging problem.

2.1. Sampling Process

A key function required by all WLAN location determination systems is signal strength sampling. For the purpose of this paper, we used a Lucent Orinoco silver NIC supporting up to 11 Mbit/s data rate.

We modified [1] the Lucent Wavelan driver for Linux so that it returns the signal strength of beacon frames received from all access points in the NIC range using active scanning [17]; our driver was the first to support this feature under Linux. We also developed a wireless API [1] that interfaces with any device driver that supports the wireless extensions [2]. The device driver and the wireless API have been available for public download and have been used by others in wireless research.

2.2. Noisy Characteristics

The IEEE 802.11b standard [17] uses radio frequencies in the license-free band at 2.4 GHz. Although this license-freedom helped in the wide spread usage of the 802.11b based networks, it also has its problems. In the 2.4GHz band, Bluetooth devices, 2.4 GHz cordless phones, microwave ovens and other devices can be a source of interference [15]. Moreover, 2.4GHz is the resonance frequency of water and human bodies can absorb the RF signal.

Multi-path fading [15, 19] is another common phenomenon in RF wave propagation. A transmitted signal can reach the receiver though different paths, each having its own amplitude and phase. These different components combine and produce a distorted version of the transmitted signal. Moreover, changes in the environmental conditions, such as temperature or humidity, affect the received signal strength.

Figure 1 gives a typical example of the normalized histogram of the signal strength received from an access point at a fixed location. People moving in the environment, doors opening and closing, and other changes in the environment can explain such temporal changes shown in the figure.

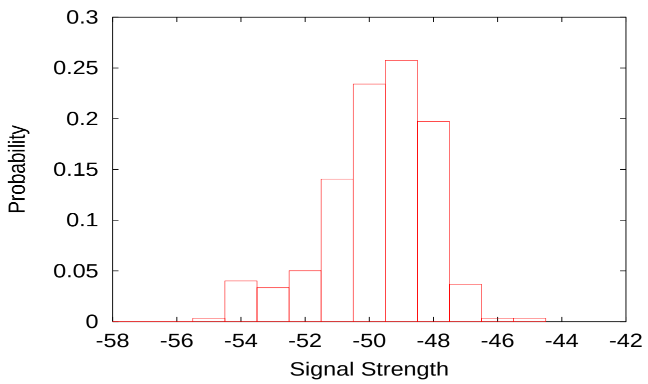

Figure 1. An example of a histogram of the signal strength of an access point

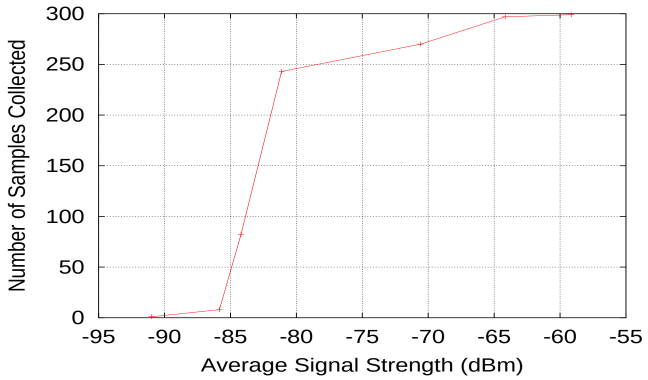

Figure 2. Relation between the average signal strength of an access point and the percentage of samples received from it.

We also performed an experiment to test the behavior of access points with different average signal strength at the same location. During this experiment, we sampled the signal strength from each access point at the rate of one sample per second. Figure 2 shows the relation between the average signal strength received from an access point and the percentage of samples we receive from it during a period of 5 minutes. The figure shows that the number of samples collected from an access point is a monotonically increasing function of the average signal strength of this access point. Assuming a constant noise level, the higher the signal strength, the higher the signal to noise ratio and the more probable it becomes that the 802.11b card will identify the existence of a packet. The sharp drop at about -81 dBm can be explained by noting that the receiver sensitivity for the card we used was -82 dBm (at 11 Mbps).

To summarize, Figures 1 and 2 highlight characteristics of the wireless channel:

- At a fixed location, the signal strength received from

an access point varies with time.
- The number of access points covering a location varies with time.

The next section presents the Joint Clustering technique that addresses these noisy characteristics.

3. The Joint Clustering Technique

We define a *cluster* as a set of locations sharing a common set of access points. We call this common set of access points the *cluster key*. The Joint Clustering technique uses the joint probability distributions of the signal strength of different access points to find the most probable user location given the observed signal strength values. Moreover, it uses clustering to reduce the computational overhead. Therefore, the operation of the Joint Clustering technique can be divided into two phases: (1) offline phase, in which we perform the joint distribution estimation and locations clustering and (2) location determination phase, in which we run the location determination technique to infer the user location. Below, we describe the two phases in more details.

3.1. Offline training phase

During the offline phase we perform two tasks: joint probability distribution estimation and location clustering.

3.1.1. Estimating the Joint Signal Strength Distribution

At each location in the set of training locations, we store a model for the joint probability distribution of the access points at this location. Therefore, our radio map is stored as a collection of models for the joint probability distributions. The problem of estimating the joint distributions can further be divided into three sub-problems:

1. How to choose a value, k, for the dimension of the joint distribution?
2. Which k access points, from the set of access points covering a certain location, to choose to be included in the joint distribution?
3. How to estimate the joint distribution between the chosen k access points?

Determining the best value for k. In determining the best value for k we need to take into account 2 factors: (1) as k increases, the process of estimating the joint probability distribution (sub-problem 3) becomes more complex and (2) we need a value for k such that all locations are covered by at least k access points most of the time. The second factor is important because the number of access points at a given location is varying with time (Figure 2). The second factor lessens the effect of variability in the number of access points and hence should lead to better accuracy. Typical values for the parameter k can be found in Section 4.

Choosing the k access points. If the number of access points covering a location is varying with time, which access points should we choose? Intuitively, we should choose the access points that appear most of the time in the samples. Figure 2 suggests that we should choose to use the k access points with the largest signal strength values at each location.

To summarize, for a given location l, we choose the k strongest access points covering this location.

Estimating the joint probability distribution. The joint probability distribution can be estimated in different ways with different accuracy levels. The problem can be stated as: given k access points $AP_1, ..., AP_k$, we want to estimate $P(AP_1 = s_1, ..., AP_k = s_k)$ where s_i is a signal strength value from AP_i. One good way to estimate the joint distribution is to use the Maximum Likelihood Estimation method which estimate the joint probabilities as:

$$P(AP_1 = s_1, ..., AP_k = s_k) = \frac{Count(s_1, s_2, ..., s_k)}{\text{Size of Training Data}} \quad (1)$$

that is, the number of times that the signal strength values tuple $(s_1, s_2, ..., s_k)$ appeared in the entire training set divided by the size of the training set.

The problem of this approach is that it requires a large training set to obtain good estimate of the joint distribution and the required size increases exponentially with k. Therefore, this approach can only be used with small values of k, which may affect the technique accuracy.

Because our goal was to use a method that gives a good accuracy and, at the same time, requires reasonable amount of training data and computational power, we choose to assume that the access points are independent. This assumption is justifiable for a well designed 802.11 network, where each access point runs on a non-overlapping channel. Moreover, independence is a common Bayesian assumption as discussed in Section 3.2. Therefore, the problem of estimating the joint probability distribution then becomes the problem of estimating the marginal probability distributions as:

$$P(AP_1 = s_1, ..., AP_k = s_k) = P(AP_1 = s_1)...P(AP_k = s_k) \quad (2)$$

For a given location, $P(AP_i = s_i)$ can be estimated using the normalized histogram of the access point AP_i at this location. Figure 1 gives a typical example of the signal strength normalized histogram from an access point.

3.1.2. Locations Clustering

To reduce the computation overhead, we group the locations into clusters according to the access points that cover the locations. The problem can be stated as follows: Given a location l, we want to determine the cluster to which l belongs.

The most obvious way to do clustering is to group locations according to the access points that cover them. i.e. two locations l_1 and l_2 are placed in the same cluster *iff* the set of access points covering these locations are identical. However, this approach for clustering has problems when applied in a real environment. As shown from Figure 2, an access point may be missing from some of the samples and, therefore, using the entire set of access points that cover a location for clustering may fail to find the correct cluster due to the missing access point.

Instead of using the set of access points that cover a location, we use a subset of this set containing only q elements and the problem becomes: Given a number q, we want to put all the locations that share q access points in one cluster. Therefore, we have 2 sub-problems:

1. How to determine the value of q?
2. Which q access points to choose for clustering?

For the first sub-problem, we need to choose q such that all locations are covered by at least q access points most of the time. This factor is important due to similar reasons as in the discussion of the choice of a value for the parameter k. This suggests that the value of q should be less than or equal to the minimum number of access points covering any location in the radio map. Moreover, we need a value for q that distributes locations evenly between the clusters to reduce the required computations. Typical values for q are given in Section 4.

For sub-problem 2, we chose to use the q access points with the largest signal strength values at each location, again for similar reasons as in the previous section.

During the data analysis we found that, at some locations, the order of the access points with the largest signal strength values changes when the signal strength values from these access points are near to each other. Therefore, we choose to treat the q access points as a set and not as an ordered tuple.

To summarize, for a given location l, we use the set of the q strongest access points covering this location to determine the cluster to which it belongs. Therefore, the cluster key is the set of the q access points used to group the locations in this cluster.

We want to emphasize here that the values of the parameters k (dimension of the joint distribution) and q (number of access points to use in clustering) are independent.

3.2. Online location determination phase

The general idea of what happens during the location determination phase is as follows: we get samples from some access points at an unknown location. We use the q strongest access points to determine one cluster to search within for the most probable location. We, then, use Baye's theorem to estimate the probability of each location within the cluster given the observed samples and the radio map built during the offline phase. The most probable location is reported as the estimated user location. The above algorithm works assuming ideal wireless channel. However, for a practical environment, we need to tackle two problems:

1. The number of access points in a test sample at a location may be less than q, the number of access points used in clustering.
2. The number of access points in a test sample at a location may be less than k, the dimension of the joint distribution.

To solve the first problem, we search for all clusters whose key has $\{AP_1, AP_2\}$ as a subset. We use the union of all the locations in these clusters as our target locations set. The set of target locations reduces to the locations within one cluster if the number of access points in a test sample is greater than or equal to q.

For the second similar problem, we use the same approach to solve it by reducing the dimension of the joint distribution to $min(k$, number of access points in the test sample$)$.

The only thing that remains to be explained is how to use Baye's theorem to calculate the most probable location out of the target locations set given the observation vector $\bar{S} = (s_1, ..., s_k)$. We want to find l such that $P(l/\bar{S})$ is maximized. i.e. we want

$$\text{argmax}_l[P(l/\bar{S})] \tag{3}$$

Using Baye's theorem, this can be rewritten as:

$$\text{argmax}_l[P(l/\bar{S})] = \text{argmax}_l[\frac{P(\bar{S}/l).P(l)}{P(\bar{S})}] \tag{4}$$

since $P(\bar{S})$ is constant for all l, we can rewrite equation 4 as:

$$\text{argmax}_l[P(l/\bar{S})] = \text{argmax}_l[P(\bar{S}/l).P(l)] \tag{5}$$

$P(l)$ can be determined from the user profile based on the fact that if the user is at a given location, it is more probable that he will be at an adjacent location in the future. If the user profile information is not known, or not used, then we can assume that all the locations are equally likely and the term $P(l)$ can be factored out from the maximization process. Equation 6 becomes:

$$\text{argmax}_l[P(l/\bar{S})] = \text{argmax}_l[P(\bar{S}/l)] \tag{6}$$

As explained in Section 3.1, the remaining term is calculated by using:

$$P(\bar{S}/l) = \prod_{i=1}^{k} P(s_i/l) \tag{7}$$

$P(s_i/l)$ is estimated from the signal strength distributions stored in the radio map

Instead of using one signal strength sample from each access point to estimate the user location, one can use a *sequence* of n samples at the same time. Assuming independence of samples, then for a given location, the probability of a sample sequence is obtained by multiplying the probability of each sample.

Using an *observation sequence*, rather than one sample, has the following advantages:

- Since we have more information from the more samples, the accuracy should be enhanced. Details in Section 4.
- This also helps in clustering: As the sequence length increases, the probability that an access point will be missing decreases. If the probability that we get a sample from an access point is p, then the probability that we get at least one sample in a sequence of n samples is $1-(1-p)^n$, . For example, if $p = 0.8$ and $n = 2$ the probability of getting a sample increases from 0.8 to 0.96.

The next section presents a discussion of the Joint Clustering technique.

3.3. Discussion

The memory requirements of the algorithm are limited. If the average number of access points per location is 4 and average range of each access point is 11 distinct values, then for each location we need to store 11*4 parameters, corresponding to the histograms of each access point, which is a small number. We could, instead, approximate the histogram by a continuous distribution, e.g. a lognormal distribution, and save only the mean and variance of the distribution for each access point. However, this approximation affects the accuracy of the system and the saving of the memory requirement does not justify it.

The clustering techniques used by the Joint Clustering technique reduces the search space and thus lead to a reduction of the computational cost. Moreover, using clustering helps in scaling the system to a larger coverage area.

4. Experimental Evaluation

In this section, we discuss the experimental testbed and evaluate the performance of the Joint Clustering technique comparing it to previous work in the area of WLAN location determination. All systems were implemented in the same environment for fair comparison.

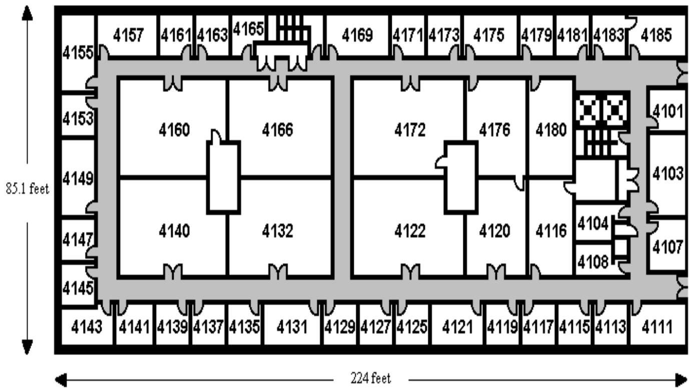

Figure 3. Plan of the south wing of the 4th floor of the Computer Science Department building.

4.1. Experimental Testbed

We performed our experiment in the south wing of the fourth floor of the Computer Science Department building. The layout of the floor is shown in Figure 3. The wing has a dimension of 224 feet by 85.1 feet. The Joint Clustering technique was tested in the Computer Science Department wireless network.

For building the radio map, we took the radio map locations on the corridors on a grid with cells placed 5 feet apart (the corridor's width is 5 feet). We have a total of 110 locations along the corridors. On the average, each location is covered by 4 access points.

Using the device driver and the API we developed, we collected 300 samples at each location, one sample per second, and we used it to estimate the distribution of each access point at each location (radio map) using the method previously described. The effect of the size of the training data on performance is discussed in Section 4.4. To test the performance of the system, we used an independent test set that was collected on different days, time of day, and by different persons than the training set.

4.2. Performance Measures

- **Accuracy**: This measure is defined as the percentage of time in which the technique gives the correct location estimate within a certain distance.
- **Number of operations per location estimate**: This measure is defined as the total number of operations (multiplications) performed for a single location estimate. This is important in minimizing computation time, but more so in minimizing the power consumption.

4.3. Effect of the Parameters on Performance

The Joint Clustering technique has two control parameters. In this section, we study the effect of these parameters, specifically k (dimension of the joint distribution) and q (number of access points to use in clustering) on its performance.

We start by showing the effect of changing q on the clustering process. For this experiment, we changed the value of q from 1 to 4 and calculated the number of clusters, the average size of each cluster, and the standard deviation of the cluster size. This is shown in Figure 4. From the figure we can see that as q increases, the number of clusters increases and the average size of each clusters decreases until we reach a saturation point at $q = 2$. For the standard deviation, the variation of the size of the clusters decreases until we reach a minimum value, at $q = 3$, and it increases again. A small value for the standard deviation means that the sizes of the clusters are more uniform, which is a desirable property. The minimum value at $q = 3$ can be explained by noting that as q increases from 1 to 3, more locations are differentiated into different clusters due to the addition of new access points. When q is increased past 3, i.e. $q = 4$, different locations start to share the same 4 access points, especially for locations close to each other (recall that the average number of access points per location was 4 in our experiment), and thus the number of locations per cluster starts to deviate from being uniform across clusters leading to increased standard deviation.

Figures 5 and 6 show the effect of parameters q and k together on performance. From the figures we see that as dimension k increases, the accuracy increases as we have more information due to the addition of access points and, due to the same reason, the number of operations required per location estimate increases. As the number of access points used in clustering (q) increases, the number of elements per cluster decreases leading to increased accuracy and less number of operations per location estimate.

For the rest of the paper, we chose to take the values of the parameters as $q = 3$ and $k = 4$ as these values lead to the best performance for the Joint Clustering algorithm for our experiment.

4.4. Results

In this section, we show the performance evaluation of the Joint Clustering technique. We also compare its performance to that of the Radar system [4], implemented in the same testbed, as a reference point.

Figure 7 shows the CDF of the error distance for both techniques. The Joint Clustering technique gives more than 90% accuracy to within 7 feet, compared to a 38% accuracy for the same distance range in the Radar system.

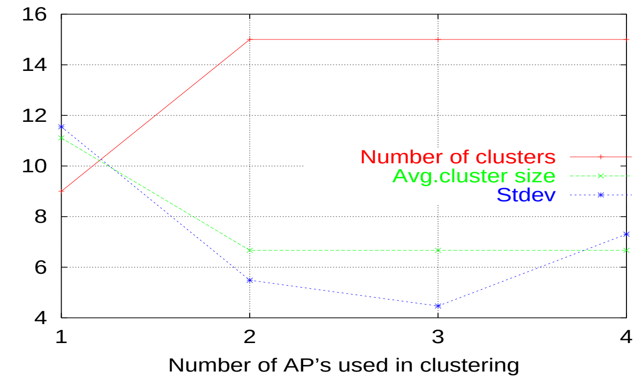

Figure 4. Effect of q on the clustering process.

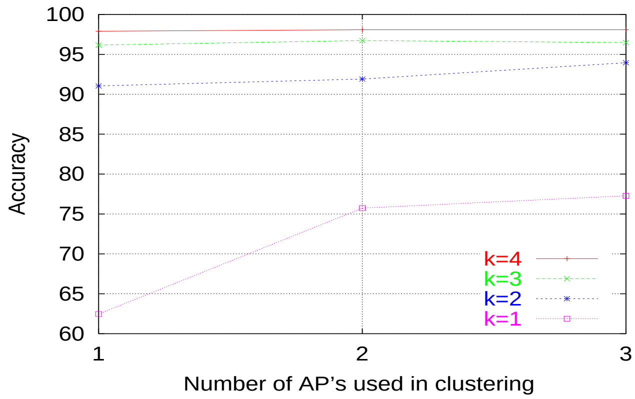

Figure 5. Effect of the parameters q and k on accuracy (within 14 feet).

A comparison between both techniques in terms of the average number of operations per sample is shown in Figure 8. The figure shows that using clustering reduces the average number of operations per location estimate by more than an order of magnitude.

4.5. Impact of the test sequence length

This section studies the effect of increasing the length of the observation sequence, used in location determination phase, on the performance of the algorithm. Figures 9 and 10 show the results. As the length of the observation sequence increases, the accuracy of both techniques increases till it reaches a saturation point at 3 samples. This is expected as the more samples we have the more information we have about the signal strength distribution and hence better the accuracy. The number of operations per location estimate increases linearly with the increase of the length of the observation sequence for the Joint Clustering technique.

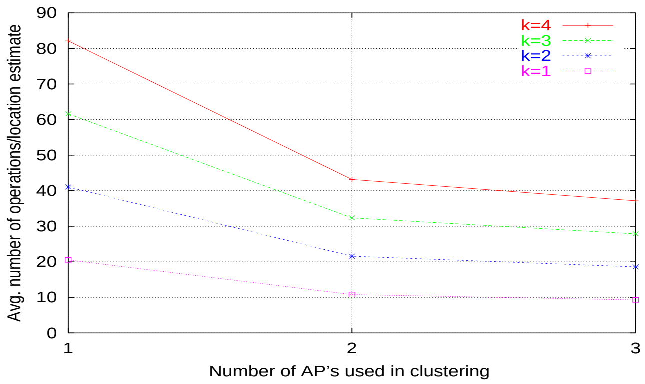

Figure 6. Effect of the parameters q and k on the avg. number of operations/location estimate.

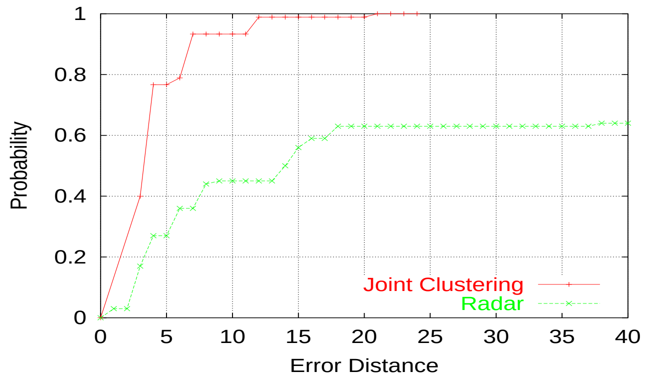

Figure 7. Error Distance CDF for both techniques (CDF for the Radar system truncated at 40 feet).

4.6. Impact of training time

Figure 11 shows the effect of changing the training time on accuracy to within 14 feet. The figure shows that the Joint Clustering technique maintains its high accuracy with a small training data set (corresponding to fraction of a minute of sampling per location).

5. Conclusions and Future Work

In this paper, we presented the design, implementation, and evaluation of a novel probabilistic indoor location determination technique: the Joint Clustering technique. The technique depends on (1) probability distributions to handle the noisy characteristics of the wireless channel, and (2) clustering to manage the computational cost.

We introduced clustering of radio map locations as an approach to reduce the complexity of the location determination algorithms and showed that clustering of radio map locations is a challenging problem with the noisy characteristic of the wireless channel. The clustering technique

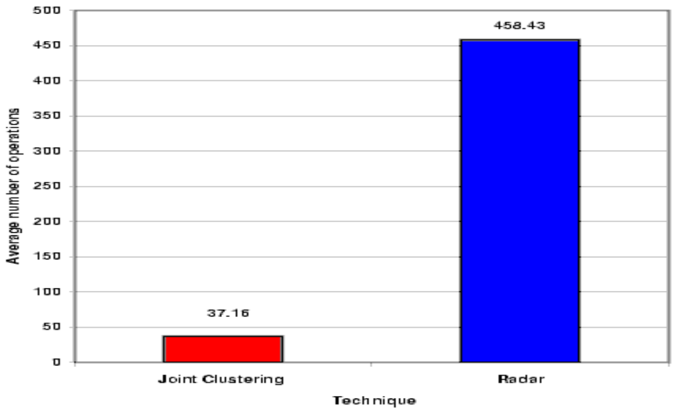

Figure 8. Computational requirements for both techniques.

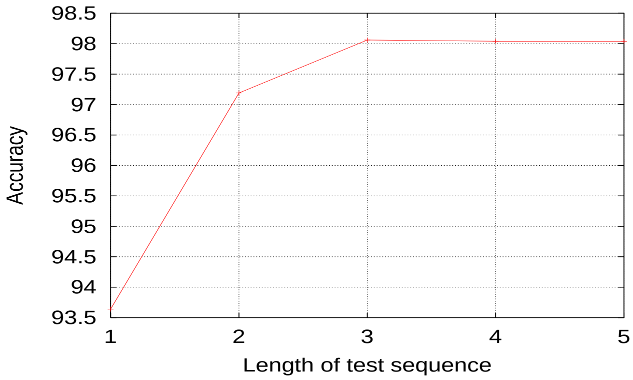

Figure 9. Impact of the length of the test sequence on accuracy (within 14 feet).

reduces the computational power by more than an order of magnitude. Such energy saving allows the system to be implemented on energy-constrained mobile devices and thus increases the scalability of the system in terms of the number of supported users. We also showed that locations clustering increases the accuracy of the location determination system and help scales the system to larger coverage area. The proposed clustering technique can be applied to all current WLAN location determination systems to reduce their computational cost and enhance their accuracy. Implementation results show that the Joint Clustering technique leads to an accuracy of more than 90% to within 7 feet.

The technique presented in this paper can be applied in both indoor and outdoor environments. Moreover, it is general enough to be applied to other technologies such as Bluetooth.

Currently we are working further to enhance accuracy and reduce computational cost. By using the user history profile and better clustering techniques, the accuracy of the location determination techniques can be enhanced. Interpolating between a number of the most probable locations is another direction that we are looking into to improve the

Figure 10. Impact of the length of the test sequence on computational requirements.

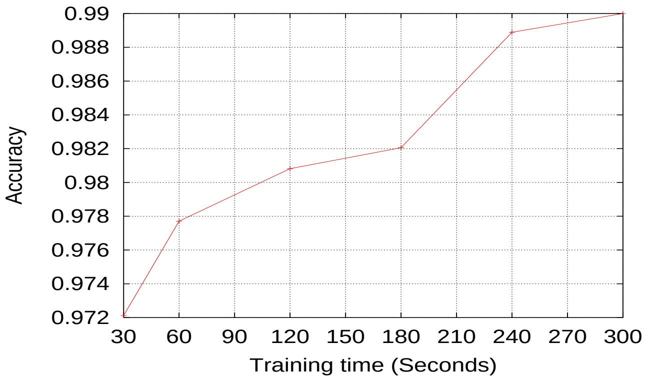

Figure 11. Impact of training time on accuracy (within 14 feet)

accuracy. We believe that understanding the nature of the radio channel and building accurate models for it are important for building more accurate location determination systems for the indoor environments and for reducing the overhead of building the radio map.

Our results gave us confidence that, despite the hostile nature of the wireless channel, we can infer the user location with a high degree of accuracy and low computational cost, hence enabling a set of context- aware applications for the indoor environments.

Acknowledgment

This work was supported in part by the Maryland Information and Network Dynamics (MIND) Laboratory, its Founding Partner Fujitsu Laboratories of America, and by the Department of Defense through a University of Maryland Institute for Advanced Computer Studies (UMIACS) contract.

References

[1] http://www.cs.umd.edu/users/moustafa/Downloads.html.

[2] http://www.hpl.hp.com/personal/Jean_Tourrilhes/.

[3] R. Azuma. Tracking requirements for augmented reality. *Communications of the ACM*, 36(7), July 1997.

[4] P. Bahl and V. N. Padmanabhan. RADAR: An In-Building RF-based User Location and Tracking System. In *IEEE Infocom 2000*, volume 2, pages 775–784, March 2000.

[5] P. Castro, P. Chiu, T. Kremenek, and R. Muntz. A Probabilistic Location Service for Wireless Network Environments. *Ubiquitous Computing 2001*, September 2001.

[6] G. Chen and D. Kotz. A Survey of Context-Aware Mobile Computing Research. Technical Report Dartmouth Computer Science Technical Report TR2000-381, 2000.

[7] P. Enge and P. Misra. Special issue on GPS: The Global Positioning System. *Proceedings of the IEEE*, pages 3–172, January 1999.

[8] A. Harter, A. Hopper, P. Steggles, A. Ward, and P. Webster. The anatomy of a context-aware application. In *5th ACM MOBICOM*, August 1999.

[9] J. Krumm et al. Multi-camera multi-person tracking for Easy Living. In *3rd IEEE Int'l Workshop on Visual Surveillance*, pages 3–10, Piscataway, NJ, 2000.

[10] A. M. Ladd, K. Bekris, A. Rudys, G. Marceau, L. E. Kavraki, and D. S. Wallach. Robotics-Based Location Sensing using Wireless Ethernet. In *8th ACM MOBICOM*, Atlanta, GA, September 2002.

[11] R. J. Orr and G. D. Abowd. The Smart Floor: A Mechanism for Natural User Identification and Tracking. In *Conference on Human Factors in Computing Systems (CHI 2000)*, pages 1–6, The Hague, Netherlands, April 2000.

[12] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan. The Cricket Location-Support system. In *6th ACM MOBICOM*, Boston, MA, August 2000.

[13] T. Roos, P. Myllymaki, H. Tirri, P. Misikangas, and J. Sievanen. A Probabilistic Approach to WLAN User Location Estimation. *International Journal of Wireless Information Networks*, 9(3), July 2002.

[14] A. Smailagic, D. P. Siewiorek, J. Anhalt, D. Kogan, and Y. Wang. Location Sensing and Privacy in a Context Aware Computing Environment. *Pervasive Computing*, 2001.

[15] W. Stallings. *Wireless Communications and Networks*. Prentice Hall, first edition, 2002.

[16] S. Tekinay. Special issue on Wireless Geolocation Systems and Services. *IEEE Communications Magazine*, April 1998.

[17] The Institute of Electrical and Electronics Engineers, Inc. IEEE Standard 802.11 - Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications. 1999.

[18] R. Want, A. Hopper, V. Falco, and J. Gibbons. The Active Badge Location System. *ACM Transactions on Information Systems*, 10(1):91–102, January 1992.

[19] M. Youssef and A. Agrawala. Small-Scale Compensation for WLAN Location Determination Systems. In *IEEE WCNC 2003*, March 2003.

[20] M. Youssef, A. Agrawala, A. U. Shankar, and S. H. Noh. A Probabilistic Clustering-Based Indoor Location Determination System. Technical Report UMIACS-TR 2002-30 and CS-TR 4350, University of Maryland, College Park, March 2002. http://www.cs.umd.edu/Library/TRs/.