Improving Accuracy of Wireless LAN based Location Determination System using Kalman Filter and Multiple Observers

Raman Kumar K, Yogesh A Powar, Varsha Apte
Dept. of Computer Science and Engineering
Indian Institute of Technology Bombay
Mumbai, Maharastra, India
Email: {raman,yogesh,varsha}@cse.iitb.ac.in

***Abstract*— RF based location determination systems are prone to problems like multipath fading and interference. These problems are not completely avoidable; but to counter them, we can employ statistical methods which are less sensitive to noise. The advantages of RF based techniques can then outweigh their disadvantages.**

We have improved the accuracy of an existing WLAN based location determination system by filtering out the noise in received signal strength readings using Kalman filter and by combining location information from more than one observer. Results from a IEEE 802.11b based implementation of our technique show that the location determination system determines the user location within 2.5 meters precision with 90% accuracy while the implementation of earlier methods achived the highest accuracy of 82% at precision of 2.5 meters.

I. Introduction

The mobile revolution has substantially changed the way people communicate. Consequently the importance of value added services which can be provided through wireless interfaces has increased manifold. The knowledge of the position of the users can facilitate a wide variety of location specific services, which could provide users with anytime, any where access to the right information at the right time. Location determination of a mobile user in outdoors is well studied and there exist many systems such as the *Global Positioning System* (GPS) [1] of USA, GLObal NAvigation Satellite System (GLONASS) of Russia and Galileo project of Europe to determine the location with accuracies upto one meter. On the other hand, the problem of locating the an indoor user is being studied only recently. The use of infrared, RF and ultrasonic technologies have been proposed for indoor location determination [2], [5], [8], [3], [6]. ActiveBadge [2] of AT&T is such a system which is based on infrared technology while ActiveBat [3] of AT&T and Cricket [6] are based on ultrasonic technology. Infrared technology is not being considered for use in large systems due to its short range visibility and requirement of line of sight. Ultrasonic based systems can locate the user with greater accuracies but are very costly. As a trade off between the accuracy of location determination system and the cost, RF based methods are also being explored.

Wireless LANs work on Radio Frequency (RF) technology which provide high data rates, upto 11 Mbps on IEEE 802.11b [4] based networks and 108 Mbps on IEEE 801.11a/g based networks. These high data rates have allowed the emergence of bandwidth-thirsty data services such as multimedia services. Traditionally Wireless LANs were only used to provide data services while using non wireless LAN methods to locate the user. However, with introduction of RADAR[5] which uses wireless LAN based RF methods to determine the location of the user, wireless LANs have grown in popularity for providing cost effective location based services indoors.

After the introduction of RADAR[5] in the year 2000, many improvements to location determination algorithm have been suggested. [8] has suggested the use of probability distribution as against the use of mean and median for storing the finger prints. Increased computational needs of this better statistic have been dealt using clustering technique in [8]. Other attempts to improve the accuracy of system have been by using user based profile information and using mobility models of users as in [16], [15], [8], [17]. Our method uses filtering of received signal strength values and employing multiple observers in location estimation, applying which the accuracy of the system is improved. Our method is independent with the methods discussed above and can be applied over these accuracy improvement methods to improve accuracy to even greater levels.

Rest of the paper is organized as follows. In section 2, we discusses about the background of developing the RF based location determination systems. We also discuss breifly the probability distribution based scene analysis method for location determination proposed by [8]. We introduce RSS filtering in section 3 and multiple observers in section 4. Section 5 discusses the filtering of location estimates got from different observers. And the whole algorithm for location determination is discussed in section 6 and the experimental setup along with results of our experiments are discussed in section 7.

II. Background

In this section we briefly discuss the organization of IEEE 802.11b infrastructure networks, their RF channel characteristics, existing RF based location determination methods and in particular scene analysis based RF location estimation methods.

A. IEEE 802.11b

An IEEE 802.11b infrastructure wireless LAN is based on cellular architecture where the system is divided into cells. Each cell, known as *Basic Service Set* **(BSS)** is controlled by an *Access Point* **(AP)**. The whole interconnected wireless LAN composing of underlying wired backbone, access points and the cells that are covered by these access points is called an *Extended Service Set* **(ESS)**. The ESS is seen as single IEEE 802 network by the upper layers of the OSI model.

Direct communication between an access point and the IEEE 802.11b based radio card on a mobile station (MS) occurs over a common channel frequency. The channel is set by the access point, and the MS's radio card tunes its transceiver to the frequency of the access point having the strongest signal. Once one channel is chosen, the mobile station associates itself to the AP and is ready to communicate with it. To support roaming, the MS's radio card periodically scans all the APs and if it finds a higher signal strength from different access point, the MS re-associates with the AP that is providing the better signal strength. The IEEE 802.11b standard defines a total of 14 frequency channels while the FCC allows only channels 1 through 11. Each of these channels, spreads over 30MHz of frequency spectrum when the total frequency spectrum available with IEEE 802.11b is approximately only 95MHz, which means, these channels overlap with each other. The only channels that do not overlap with each other are 1, 6 and 11 which can be used without causing interference between access points.

B. Characteristics of RF channels

The decrease in strength of the signal when transmitted over a distance is called the attenuation. The *received signal strength index* **(RSSI)** of an AP at a location is an indicator of how strong the signal from that access point is at that location. RSSI is measured in dBm as shown in Equation 1 and is used in the process of mapping the location and estimation of a mobile user's location.

$$RSSI\ [dBm] = 10\log\left(\frac{received\ power\ in\ watt}{0.001\ watt}\right) \qquad (1)$$

RF channels are prone to noise, and when used indoors the effect of the noise on RSSI values is much more than when used outdoors. Below is the list of noise factors in IEEE 802.11b channels:

- Multipath Fading [7]
- Interference
- Presence of human and water
- Change in temperature and moisture

With the presence of noise, it has been observed by [8] that:

- At a fixed location, RSSI value of the Access Points measured at the MS is not stable; while being distributed within a small interval.
- The number of access points covering a location varies with time.

It is also observered that even with the presence of these noise factors, RSSI can be used as an feature particular to a location. This is because of the greater signal to noise ratio (SNR).

C. Existing RF Methods

Proximity, triangulation and scene analysis are the three broad classes of RF location estimation methods. Proximity and scene analysis methods are explained below.

1) WLAN based Proximity: In this method, the area covered by the wireless network is divided into non overlapping regions; one region for each access point. The location is classified to belong $region_i$ if access point AP_i dominates in terms of RSSI value at that location for most of time. Also, the MS gets binds itself to the AP which gives the highest RSSI. Depending on the access point to which the MS binds, the location determination system estimates the location of the MS to the region covered by that particular access point.

2) WLAN based Scene Analysis: In this method, the received signal strength index is measured at the MS's Wireless LAN adapter card and is used as the finger print measurement of that location. All the scene analysis methods are two phase methods, having the offline training phase and the online estimation phase. In offline phase the location determination system is trained about various locations and their corresponding RSSI finger print measurement. And in online phase, location is determined by the finger print extracted at that particular location.

D. WLAN based Scene Analysis

In this section we discuss the widely used scene analysis methods.

1) Simple Mean Method: In this method a single finger print is saved in the radio-map as an identifier of location. During the training phase, we sample large number of measurements, N, typically in order of tens at all the locations with predetermined granularity. Let the i^{th} reading be represented by $< rss_{1i}, rss_{2i}, ..., rss_{ni} >$, where rss_{ji} being the received signal strength from access point AP_j. Once the samples are collected the representative is calculated as $< \Sigma rss_{1i}/N, \Sigma rss_{2i}/N, \Sigma rss_{3i}/N >$ and is stored in the radio-map as a finger print of that particular location.

In the online phase, the RSSI feature of the location is measured at the required location, which is then compared with the RSSI finger prints of all sampled locations. The location which has closest RSSI finger print in terms of Euclidean distance is reported back as the identified location. This method is simple to implement and is computationally effective but does not deal with the problems caused by access points whose beacons are missed in greater numbers. As a statistical measurement, mean does not incorporate this data. Also this way we will not be haveing information of the variance of the distribution of RSSI of a particular access point at the location.

2) Bayseain Method: Considering the problems with using mean, use of probability distribution for storing finger prints of various locations is introduced in [8]. In this method, the number of RSSI measurment tuples are read and stored in the radio map in the form of probability distribution. And in the online phase, RSSI feature at the location is extracted and this measurement tuple is used to query the location determination system. Location determination system reports the location which have the highest posterior probability of observing the extracted tuple.

III. Noise Filtering in RSSI measurements

At a particular location, the signal strength received should only depend on the attenuation factor of the media between the mobile station and the access point; but during the measurement of this signal strength, various factors such as the interference by other devices, presence of human and water and multipath fading add a noise component to the signal measured. In addition to the above measurement noise, there is another type of noise within the system, noise in the power transmitted by the access point. Also, the power transmitted by the access point is not modified with any control inputs. We can observe that the whole scenario fits into the scenario for which Kalman filter [13], [9] equations can be used. State x that is being measured in our case is the RSSI of a particular access point at a location. Measurement of the RSSI is effected by the measurement noise produced by other devices operating at the same frequency. The noise in the access point which effects the transmission power is the process noise. Process noise in the access point is close to zero and negligible when compared to the measurement noise. The measurement noise variance is estimated to be 4 units. Kalman filter equations does not need this noise variances to be exact, the value of these variances converge when large number of samples are taken. Therefore, our initial estimation is the noise will only effect the number of samples needed. In our case we take 20 samples for RSSI filtering.

IV. Multiple Observers

All the location determination methods discussed till now involve only one perspective, the perspective the mobile station gets by scanning the beacons arrived from the access points. The mobile station calculates its location based on this perspective it gets. We conjecture that it could be beneficial to get multiple perspectives and use this extra information to improve the accuracy of the location determination system. The extra perspective the system could get is the RSSI values of the mobile station's signal measured at all the access points. This extra perspective does come with the cost that the mobile device should bind itself to the access point to get the access point measurement of mobile station signal strength. This binding will be costly in terms of time as the mobile station has to connect to all the access points which support the measurement of MS's RSSI. In our case, the multiple perspectives we get are:

A. MS Perspective

In this perspective, the mobile station scans the channels and measures the RSSI values of all the visible access points. That is, the finger print of a particular location, in this perspective is the tuple of RSSI values of signals from different access points represented by $< rss^{MS}_{AP1}, rss^{MS}_{AP2}, ..., rss^{MS}_{APk} >$. Where rss^{MS}_{APi} represents the RSS value of access point i measured at mobile station.

B. AP Perspective

In this perspective, the mobile station scans the channels to find all the visible access points. It then binds itself to each of the access point and queries about the RSSI measurement of the signal received from the Mobile station. That is, the finger print of a particular location in this perspective is the tuple of RSSI values of signals from mobile station received at different access points. The finger print is represented by $< rss^{AP1}_{MS}, rss^{AP2}_{MS}, ..., rss^{APk}_{MS} >$ where rss^{APi}_{MS} represents the signal strength of mobile station measured at the access point i.

V. Noise Filtering in Location estimates

Each of the perspectives will be estimating the location of the mobile user. The simplest way to combine the location information is to calculate the mean of these two location estimates and report the mean to the mobile user as the location. Reporting the mean will imply that we are trusting both the perspective equally. But this may not be the case, we may want to trust location reported by one perspective more than the location reported by the others. This whole scenario could also be seen as a method to estimate the location of the mobile station given that we have multiple measurements. Each of these measurements is with some noise variance. Greater the noise variance lesser we would want to believe the measurement. We apply Kalman filter to these location measurements and report the the filtered location estimate to the mobile user. In our location determination system, MS perspective and AP perspective will report location of the the mobile station independently. The location is reported as 3-D coordinate. We assume that the coordinates of the location are mutually independent as we do not have the prior knowledge of the area covered by the location determination system. Kalman filter is applied independently to each of the dimensions. The location determination system reports the filtered x-dimension, y-dimension and z-dimension as the location of the mobile station.

VI. Algorithm

As the case with any scene analysis method, our algorithm is also divided into two phases, offline and online. In the offline phase the radio-maps for both the perspectives are built. In the online phase the built radio-maps are used to answer the location queries from the mobile station.

A. Offline - Training Phase

At the mobile station, all the IEEE 802.11b channels 1 through 14 are scanned to get the list of the active access points. Beacons from these access points are used to measure the signal strength of that particular access point. Wlanctl-ng software is used to interface with the Wireless LAN adapter card. The sampling procedure is repeated 20 times to get the RSSI feature tuples at the location. The collected RSSI feature tuples are then filtered for noise and thus obtained noise reduced RSSI tuple is used as the finger print information. This finger print information is saved in the radio-map contributing to the RSSI distributions of access points. As the samples collected were at the mobile station, MS-perspective radio-map is used. The whole procedure of scanning, sampling, filtering and storing is repeated 20 times and thus obtained RSSI finger print information is used to build the RSSI distribution of all visible APs at a particular location. For the case of AP perspective, access points are classified into 2 types; type 1 are those access points that can measure the strength of the MS's signal and type 2 being the rest of the access points. All the access points that are active at the MS's location and are of type 1 are queried to report the signal strength of the MS. These readings are sampled over 10 measurements and filtered using Kalman filter. 20 such filtered measurements are then stored in the AP-perspective radio-map as RSSI distribution.

In the offline phase, we also build a list of locations for every access points, which are known to be near that particular access point. A location is said to be near a particular access point if that location receives a signal strength greater than *proximity_threshold*. We set the value of proximity_threshold to -90dBm. This *proximity list* is used in the online phase in location searching.

The algorithm of offline phase is given in Figure 1. This algorithm is used separately for MS perspective and AP perspective independently to build respective radio-maps. It can be observed that the two perspectives differ only in the way the finger prints are collected at a location.

1: Collect RSSI finger prints in this perspective at a location
2: Filter them using Kalman Filter
3: Save the filtered RSSI tuple in Radio-Map
4: If the RSSI of an AP in the filtered tuple is greater than -90dBm, add the location to *proximity list* of the AP
5: Repeat the procedure for all the location to be sampled

Fig. 1. Training the Perspective

The location information used in our algorithm is also different from the location information used in previous algorithms. Symbolic locations were sufficient for earlier algorithms which is not the case for our algorithm. We need the physical location represented by the (x,y,z) co-ordinates. As we are applying the Kalman filter for location filtering we need this physical location information.

B. Online - Estimation Phase

In online phase, the mobile station collects multiple RSSI feature tuples and filters out the noise in the measurements. Thus obtained noise free finger print of the location is used to query the radio-map of MS-perspective for its location. Similarly, RSSI feature tuples are collected for AP-perspective. These tuples are filtered and then used to query the radio-map of AP-perspective. Each of the perspective will input the finger print and output the location estimate. The location estimate is a probability distribution of observing various locations. These distributions are represented by $< mean, \ variance >$ pairs.

The per-perspective location estimation algorithm is given in Figure 2. Given the finger print whose corresponding location has to be estimated, we can find the approximate location of the mobile station by using the finger print tuple. We can say that the location of the mobile station is in the vicinity of the access point which provides the highest signal. That is this location exists in the proximity lists of the access point providing the highest signal. This exercise reduces the number of locations to be searched for most appropriate location.

1: using Proximity Lists, narrow down the search of most probable location
2: use Bayesian method to find the most probable location from above list
3: while probability of the *most probable location* at the present precision level is less than *minimum threshold probability*, degrade the precision and search again
4: Thus obtained location is reported by this perspective

Fig. 2. Online phase of the Perspective

Once the list of locations that are to be searched is narrowed down, the probability of observed finger print belonging to various locations in this list are calculated using the Bayesian method discussed. If the probability of the most probable location is itself very low, the confidence in the location determination system will be low. To increase the confidence we degrade the precision of the location determination algorithm and we search through the narrowed down list of locations again. The degradation of the precision is done iteratively until the probability of the most probable location crosses the pre decided *probability_threshold*. We have this threshold set at $(\frac{3}{4})^n$, where n is the number of visible access points at the location of the mobile user. Once such a location is found, this location is reported to the mobile station. It can be observed that degrading the precision and improving the accuracy implies that the location determination system have more confidence that the mobile station is present near to the location estimated. The degree of nearness is proportional to the precision. When we degrade the precision the size of area in which location determination system have confidence also grows. Therefore, degrading the precision increases the variance of the location estimate. We empirically calculated that with every level of degradation in the precision increases the variance of the location estimate by 1 meter. This is the heuristic we use in calculating the variance of the location estimate. Searching the location at different precision levels is given below.

We extended the location determination system proposed by [8] to degrade precision in order to improve accuracy. That is we made the precision of our location determination system to be adjusted depending upon the need of the user's application. We degrade the precision of the system by using the probabilities for a larger RSSI interval instead of selecting the precise value in calculating the joint probabilities of observed RSSI feature belonging to a particular location. Let the RSSI reading in the online phase at an unknown location be given by tuple $< rss_1, rss_2, rss_3, ..., rss_n >$. To find the location that best represents the MS's location with *precision level* a:

$$argmax_L(\prod_{i \in visible\ APs} P_{RSS_{Li}}\,(rss_i - a \geq x \geq rss_i + a)) \quad (2)$$

where $Probability_{RSSI_{iL}}$ is the distribution at location L of Access point i. *Visible* access points are those access points, whose beacons can be identified by the mobile station at this particular location.

Thus obtained location information by the MS-perspective and AP-perspective is then combined using the method discussed in Section V. The location estimated by the combined perspectives is then reported back to the mobile station as the location determined.

VII. Experimental Setup and Results

In this section, we discuss implementation of our algorithm. We also discuss our location determination algorithm's accuracy and computational cost. For the experiments, RSSI feature is measured at different locations using DWL122 USB network adapter which is interfaced using *wlanctl-ng*. The network adapter is programmed to scan channels 1 through 11 to read the beacons transmitted by access points. The network adapter scans the channel with minimum channel duration of 200 milli seconds and maximum channel duration of 250 milli seconds. These values were chosen, assuming that the access points are configured with beacon interval of less than or equal to 200 milli seconds. We used physical locations in this setup. Physical locations are measured with a reference co-ordinate system with in the building.

A. Wireless LAN Testbed

Our experimental setup is equipped with 31 access points covering the building of four floors. All of these access points are fixed, 28 of which were configured to operate in channel 6 of IEEE 802.11b standard and the rest three were configured to be in channels 1 and 11 of IEEE 802.11b standard. In this setup, there are many wireless network users, whose presence on wireless network added noise to the RF signals. Also the access points themselves will interfere with each others transmissions as most of them are configured in same channel. We had assumed that the beacon interval of all access points are configured to 200 milli seconds. For the sampling purpose we used the granularity of 1 square meter.

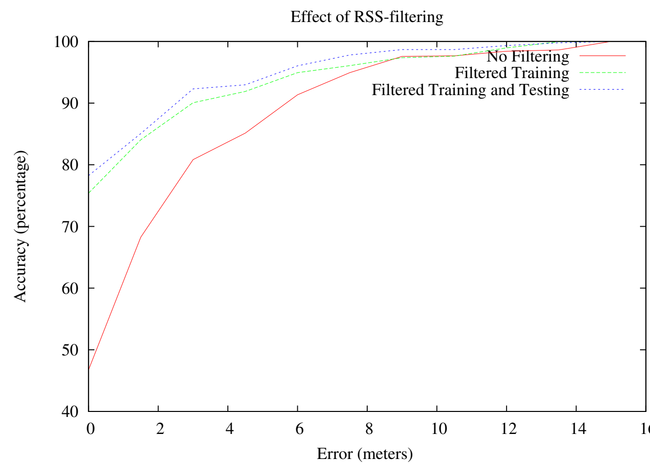

Fig. 3. Effect of noise filtering in RSS on Accuracy

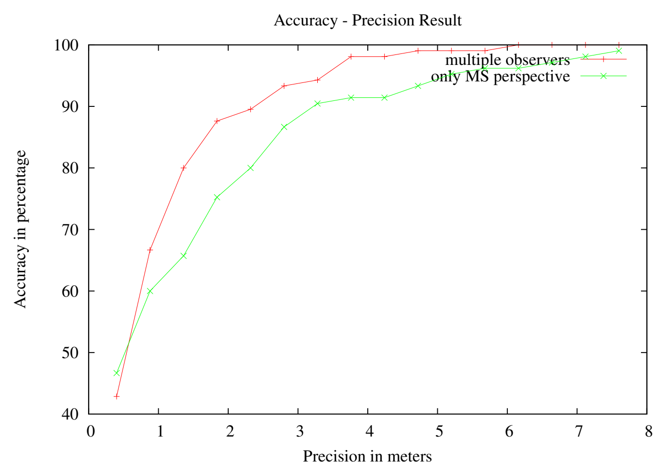

Fig. 4. Accuracy of our location determination system at various precision levels

B. Results

During the training phase, we sampled in selected locations within the area covered by the system, and we tested over location determination system with more than 400 location identification requests from different locations within the area covered by the system. The test cases also include the locations which are not sampled during the training phase which should be reported as close to their sampled neighbors by the perspectives. Distance between the actual MS's position and the location determined by the system is measured as error distance. Figure 4 shows the percentage of test cases that were reported with different error distances.

We have got the accuracy of 90% with precision of 2.5 meters. The accuracy of our location determination system at various precision levels in comparison with the method developed at University of Maryland [8]. In Figure 3, the comparitive improvements gained by appling noise filter for RSS

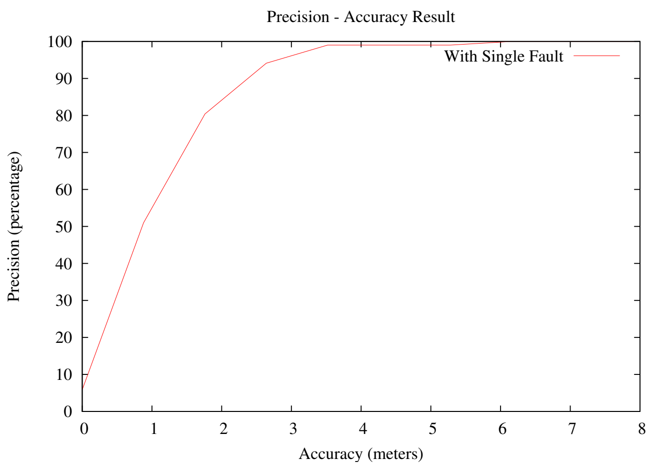

Fig. 5. Accuracy of our location determination system at various precision levels in the presence of fault

is shown. Figure 4 shows the comparision of two scenarios; one involving simple MS perspective and second involving two-perspectives with location filtering.

The other performance metric of a location determination system is the computational cost or delay in location reporting. After the location determination system gets the RSSI feature measured at the MS's location, it has to search through all the sampled locations to find the most probable location. We reduced the search space by using proximity method; using which we search only through the proximity of the most active AP (the AP whose RSSI is the strongest) as reported by the MS. That is we use proximity technique to first locate the user to the size of the building. Then we search through the locations in this building to find the most probable location. Each access point typically covers 200 locations when the sampling granularity used is 1 square meter. Therefore, the time the location determination system takes to search the most probable location is independent of the total number of locations sampled. Hence, the time complexity of our location search algorithm is O(1); independent of the area covered by the location determination system.

The location determination system will be reporting the location of the MS's given its measured RSSI feature within order of milliseconds. But the response time, the user perceive, will be in order of seconds. This is due to the delay at the mobile station in collecting the RSSI feature from all the access points; as it require to scan all the IEEE 802.11 channels. To reduce this, we need to configure the access points with much shorter beacon interval so that MS could scan through the channels fast. This shortening of beacon interval will result in flooding of the channels with beacon signals, causing greater interference between these signals and degrading the quality of the location determination system.

We had conducted an experiment with single fault, single access point missing, the result of the experiment is shown in the Figure 5. It can be observed that the system is reporting with similar accuracy levels as in the experiment conducted with out any fault, this can be attributed to the fact that there are more than 3 access points that are used for measurements even when 1 access point is faulty. The only difference being the variance with which the location determination system is reporting, variance reported in later experiment is much greater than the variance reported in the experiment with out any fault.

References

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

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

[3] Ward Andy, Jones Alan, and Hopper Andy. *A new location technique for the active office*. In *IEEE Personal Communications*, volume 4, pages 42–47. IEEE, 1997.

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

[5] Paramvir Bahl and Venkata N. Padmanabhan, *RADAR: An In-Building RF-Based User Location and Tracking System*, INFOCOM (2),Pages: 775-784, 2000.

[6] Nissanka B. Priyantha, Anit Chakraborty, and Hari Balakrishnan. *The Cricket location-support system*, In *Proc. of MOBICOM*, pages 32–43, 2000.

[7] H. Hashemi, *The Indoor Radio Propagation Channel*, In *Proc. of IEEE*, 81:943–968, 1993.

[8] Moustafa Youssef, Ashok Agrawala, and Udaya Shankar. *WLAN Location Determination via Clustering and Probability Distributions*, In *IEEE PerCom 2003*, March 2003.

[9] Peter S. Maybeck. *Stochastic models, estimation, and control*, volume 141 of *Mathematics in Science and Engineering*, Academic Press, 1979.

[10] F. Espinoza, F. Peterson, P. Sandin, H. Nyström, E. Cacciatore, and M. Bylund. *Geonotes: Social and navigational aspects of location-based information systems*, *Ubicom 2001: Ubiquitous Computing*, pages 2–17. Springer, September 30 - October 2 2001.

[11] William G. Griswold, Patricia Shanahan, Steven W. Brown, Robert Boyer, Matt Ratto, R. Benjamin Shapiro, and Tan Minh Truong. *ActiveCampus: Experiments in Community-Oriented Ubiquitous Computing*, *Computer*, 37(10):73–81, 2004.

[12] S. Koo, C. Rosenberg, H. Chan, and Y. Lee. *Location-based e-campus web services: From design to deployment*, In *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, PerCom* , 2003.

[13] R. E. Kalman. *A new approach to linear filtering and prediction problems*, In *Transactions of the ASME Journal of Basic Engineering*, pages 34–45, 1960.

[14] Aittola M, Ryhanen T, and Ojala T. *Smartlibrary - location-aware mobile library service*, In *Proceedings of Fifth International Symposium on Human Computer Interation with Mobile devices and Services, Udine, Italy*, pages 411–416, 2003.

[15] Siddhartha Saha, Kamalika Chaudhuri, Dheeraj Sanghi, and Pravin Bhagwat. *Location Determination of a Mobile Device Using IEEE 802.11b Access Point Signals*. In *IEEE Wireless Communications and Networking Conference*, 2003.

[16] Tadhg P. Deasy and William G. Scanlon. *Stepwise Refinement Algorithms for Prediction of User Location Using Receive Signal Strength Indication in Infrastructure WLANs*. *IEEE High Frequency Postgraduate Student Colloquium*, pages 116–119, 2003.

[17] Venkata N. Padmanabhan. *LandLoc: Landmark-based User Location*. Technical Report MSR-TR-2001-23, Microsoft, January 2001.