Location Estimation in ZigBee Network Based on Fingerprinting

Qingming Yao, Fei-Yue Wang, *Fellow, IEEE*, Hui Gao, Kunfeng Wang and Hongxia Zhao

Abstract— **Location-aware computing becomes an exciting research as recent advancements in RF circuits and wireless communication stacks. In this paper, we present a fingerprinting based location estimation technology in ZigBee network. The system uses the signal strength from several base stations rather than time or angle for determining the location of mobile station. Instead of modeling the complex attenuation of signal strength, the system models the probabilistic distribution in different geographical areas which we called fingerprinting. It combines the measured data and fingerprinting to determine the mobile station's location. The experiment results demonstrate the validity of location estimation in ZigBee network based on fingerprinting.**

Index Terms— **-Location Estimation; ZigBee Network; Signal Strength; Fingerprinting; Probabilistic Distribution.**

I. INTRODUCTION

Advancements in electronics, computer and wireless communication technology have promoted the development of low-power, multi-function sensors, and enabled integrating A/D converter, data processor and wireless communication modules into a single chip. On one hand, seamless communications among such devices and possible processing centers can transform ordinary environments into intelligent spaces [1]; on the other hand, the distinction between communications and computation is blurring [2]. We are being carried into pervasive communicating and pervasive computing space in which context is an obvious attribute. Context refers to the physical (position, time, weather) and social situation (work or leisure place) in which computational devices are embedded [2]. As location becomes one of the most import contexts, location-aware computing is a recent interesting research area. It can provide services such as personal security, children track, tourist guide and entertainment.

The existed location systems, including the Global Positioning System (GPS) and wireless enhanced 911(E-911) and cellular network-based usually work outdoors with coarse granularity [3], worse in indoor and building-dense environments. Another shortage is that we can't transport sensor data about the other contexts in these networks. Other systems were developed for short range, indoor and precious location using infrared, ultrasound or RF and ultrasound combined techniques. *AT&T's Active Badge*, which uses diffuse infrared technology has difficulty with fluorescent lighting or direct sunlight, but can get room-size accuracy. *Active Bats*, which also comes from *AT&T*uses an ultrasound time-of-flight lateration technique getting more accurate (9 cm), but requires large scale deployment and high cost. *MIT* complemented the Active Bats by using a radio frequency control signal, which names *Cricket*. Although it does not require a grid of ceiling sensors, the *Cricket* lacks of centralized management, and the mobile receivers have heavy computational and power burden [4]. Another critical disadvantage of all the three systems is that they only provide light-of-sight (LOS) location estimation.

As we know the 802.11 (predominant are 2.4 GHz 802.11b and 802.11g) Wireless Local Area Network (WLAN, also called Wi-Fi in business) is installed popularly, and using this data network to support ubiquitous-covered, accurate location estimation is increasingly gaining passion since Microsoft developed *RADAR*. By recording the received signal strength (RSS) from existed several 802.11 Access Points and comparing to empirical measured data, *RADAR* is able to estimate user location in office or home environment easily [5]. *Carnegie Mellon University* [6], *IBM* [7], *Maryland University* [8] and *Pittsburgh University* [9] also have started deeper research and system evaluation. These systems have almost the same sensing devices, reference points but different algorithms and infrastructures. The performances reported are similar (4 feet [6], 6 feet [7], 7 feet [8] and 9 feet [9] respectively over 90%).

All the location estimation approaches that based on WLAN's Received Signal Strength Indication (RSSI) are used in indoor environments and passively up to now. As the data network is not transparent and sensing-oriented, the context-awared services provided by such systems are limited. And the infrastructures of the network are not established for location estimation generally, in another word, the deployment, change and movement of the network are out of the location system's control.

In this paper, we implement the location estimation system adopting ZigBee based network, which is a short rage, low data rate, low power consumption and low cost network technology. ZigBee' target aims at automation and remote control applications through easily constructing ad-hoc, mesh networks, and it will provide more ubiquitous coverage. Theoretically speaking, the location estimation algorithm used in ZigBee network may be the same as used in WLAN.

The remainder of this paper is organized as follows: Section II describes the research methodology in detail of hardware, architecture and algorithm. Section III discusses implement process and analyses the results. And Section IV concludes the paper and gives summary of future work being

This work was supported in part by the MOST Grant 2006CB705500, NNSFC Grants 60334020, 60621001, CAS Grant 2F05N01, 2006 CAS Presidential Special Scholarship Grant, and SDAS Research Grant 2005006..

Qingming Yao, Fei-Yue Wang, Hui Gao, Kunfeng Wang and Hongxia Zhao are with the Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing, China. qingming.yao@ia.ac.cn

1-4244-1266-8/07/$25.00 ©2007 IEEE.

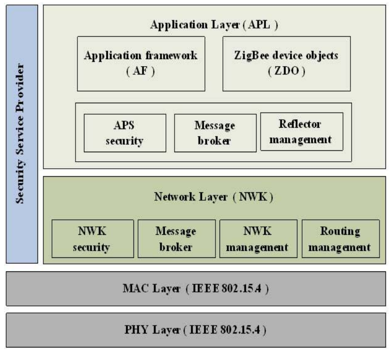

Fig. 1. Diagram of ZigBee stack.

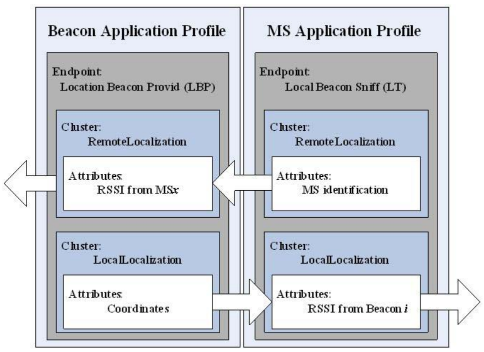

Fig. 2. Application profiles.

undertaken.

II. Research Methodology

A. ZigBee Network

ZigBee uses IEEE 802.15.4 standard as its PHY and MAC layer standard which defines a 250k bps direct sequence spread spectrum (DSSS) radio operating in the 2.4 GHz unlicensed band, with lower bit-rate alternatives in the 868 MHz and 900 MHz bands [11]. Above the MAC layer, ZigBee defined network layer (NMK) and application layer (APL) [10]-[12]. The diagram of the stack is showed in Fig.1. [12], [13].

RSSI is the basic function we use to form fingerprinting and measure data. As the multi-path affection in upstream and downstream is usually different, the RSSs received by Beacon and MS are different. Although we have defined two Clusters in Location endpoint: the *RemoteLocation* and *LocalLocation*. But in this paper, only the *LocalLocation* cluster is used. Fig.2 presents the attributes and clusters defined in the application profiles.

B. System Methodology

In our system, some Beacons are fixed in several points evenly to make sure that mobile station (MS) can receive *n* (3 to 5 as usual) points' radio signal at each location. The MS records and processes the RSS vector and then searches the fingerprinting database to find some fingerprinting which makes the algorithm criterion maximum. Each fingerprinting is formed by a phase of training in a location. Error distance is used to evaluate the result. System methodology is showed in Fig.3. $O_{xy} = (o_{xy}^1, o_{xy}^2, ..., o_{xy}^n)^T$ is the observed RSS vector from beacons at location L_{xy}, and $F_{ij} = (f_{ij}^1, f_{ij}^2, ..., f_{ij}^n)^T$ is average RSS of location Fingerprinting database is constructed by a process of offline training. Denote a series offline training measurement of beacon k at location L_{ij} is $L = [l_{ij}^{k_0}, ..., l_{ij}^{k_{M-1}}]$ which enables computing the histogram h_{ij}^k of signal strengths for each beacon indexed k:

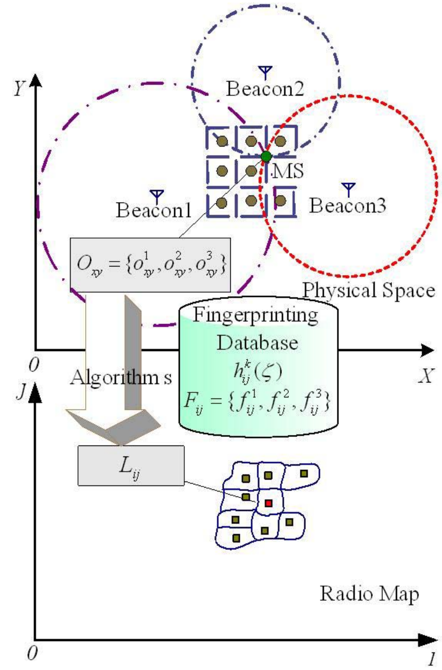

Fig. 3. System methodology (n=3).

$$h_{ij}^k(\zeta) = \frac{1}{M} \sum_{m=0}^{M-1} \delta(l_{ij}^{km} - \zeta), -255 \leq \zeta \leq 0 \qquad (1)$$

where ζ represents the Kronecker delta function [14].

The estimation algorithms will map the online observed data O_{xy} to some physical L_{xy} by using probabilistic method; detail is described in II.C. We define two error distances to evaluate the accuracy from signal and physical space respectively. The signal space's *Euclidian distance* between O_{xy} and L_{xy} is [9]:

$$Dist(O_{xy}, F_{ij}) = \sqrt{\sum_{k=1}^{n} (o_{xy}^k - f_{ij}^k)^2} \qquad (2)$$

And the physical space's *Euclidian distance* is:

$$Dist(L_{xy}, L_{ij}) = \sqrt{(x-i)^2 + (y-j)^2} \qquad (3)$$

C. Location Estimation Algorithm

Radio waves can be modeled accurately in free space [7]. But in practical environment, the radio channel is of noisy characteristics by reflection, absorption, diffraction, scatteration and even the fluctuation of temperature. So the signal transmitted from beacon usually reaches the mobile station by more than one path, resulting in a phenomenon known as mutli-path fading [13] [15] [19]. The observation O_{xy} deviates significantly from F_{ij} Fortunately, through observation the signal strength's temporal distribution presented in III.C, we conclude that the distribution has obvious statistical character. So we applied a probabilistic approach using Bayesian inference.

Suppose there are 4 beacons deployed in system. The estimation algorithm's target is to find a location L_{ij} hat makes the probability $P(L_{ij}|O_{xy})$ maximized.

Mathematically, the probability $P(L_{ij}|O_{xy})$ can be represented as:

$$\begin{aligned} P(L_{ij}|O_{xy}) &= \frac{P(O_{xy}|L_{ij})P(L_{ij})}{P(O_{xy})} \\ &= \frac{P(O_{xy}|L_{ij})P(L_{ij})}{\sum_{j=1}^{J}\sum_{i=1}^{I} P(O_{xy}|L_{ij})P(L_{ij})} \end{aligned} \tag{4}$$

where conditional probability $P(O_{xy}|L_{ij})$ is the likelihood of O_{xy} occurring in the training phase of L_{ij}. And $P(L_{ij})$ is the prior probability of location L_{ij} being the correct position [15] and is usually uniformly distributed if the position relative to the map is entirely unknown.

The CSMA/CA mechanism ensures the signal from different beacons independent from each other, so the joint probability distribution then becomes the problem of estimating the marginal probability distributions as:

$$P(O_{xy}|L_{ij}) = \prod_{k=1}^{n} P(o_{xy}^{k}|L_{ij}) \tag{5}$$

where

$$P(o_{xy}^{k}|L_{ij}) = h_{ij}^{k}(\zeta) \tag{6}$$

is the location fingerprint can be obtained in the offline training phase [14]. We can conclude two points about whether the methodology works:

1) Feasibility: $h_{ij}^{k}(\zeta)$ itself being stable enough and $\prod_{k=1}^{n} h_{ij}^{k}(\zeta)$being sensitive to different (i, j) are necessary conditions.

2) Practicability : Each histogram $h_{ij}^{k}(\zeta)$'s distribution is narrow. In another word, the standard deviation should be small. Only several ζ are high-frequent, so the small search space of the algorithm insures the practicability.

III. IMPLEMENT AND RESULTS

A. ZigBee Module

We implement a ZigBee module adopted TI's single-chip 2.4 GHz IEEE 802.14.5 compliant RF transceiver CC2420 and MicroChip's enhanced Flash and nanoWatt technology microcontroller PIC18LF4620. A LCD module was added to coordinator node for displaying some useful network information such as network finding, orphan notification, association and disassociation. And an EIA RS-232 interface is also supported to receive command from or send network information to an attached laptop. The module is presented in Fig.4 and the system components are presented in Fig.5.

For location estimation application, the wireless modules are fixed on ceiling usually. To obtain even covering in

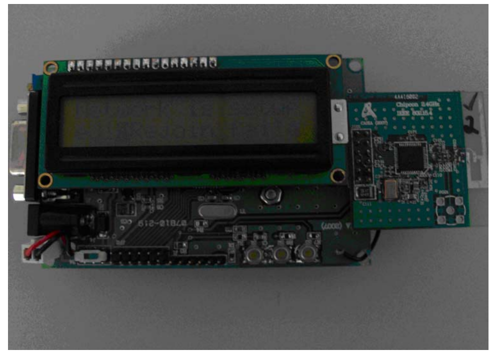

Fig. 4. Wireless module.

Fig. 5. System components.

estimation area, we apply the 2.4 GHz 50 ohm inverted-F antenna which gets 1.1 dB gain and omni-directional radiation pattern in PCB plane [22]. The module can evenly cover about 30 meters in clear space and 18 meters in indoor space as be programmed to 0 dBm output power with the typical -95 dBm receive sensitivity.

B. Layout of Experimentation

The experiments were carried out in an office room dimensions of $7.2m \times 9m \times 2.6m$. The layout was depicted in Fig. 6. All the beacons are fixed on the ceiling and powered by batteries. They are separated near the four corners that enable them provide even four-overlap coverage in all portions of the office. The calibration points where the mobile stations' signal strength was collected are denoted by the gray square. And the arrow's direction indicates the operator's orientation as the attenuation by operator's body can effect the signal strength significantly [5] [7] [9]. For carrying out experiment, we only defined 34 calibration points, and in each point we collect the signal strength only in one direction. The laptop connected with a wireless module by serial cable was put on a small desk with 0.7 meter height which can be moved stably.

C. Signal Statistical Character

Multi-path fading and people's activities lead the RSSI luctuating, such as Fig. 7 shows. We took numerous measurements at various locations under different scenes to see whether the feasibility and practicability can be obtained

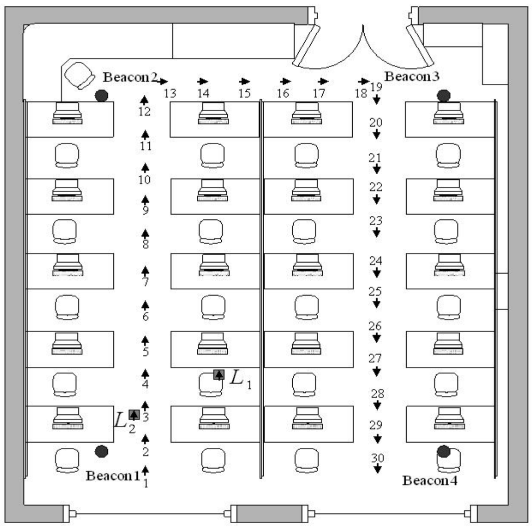

Fig. 6. Layout of the experimentation environment.

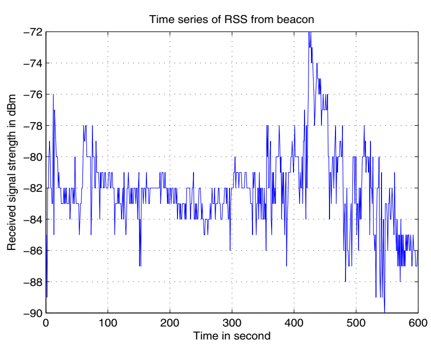

Fig. 7. Typical signal strength received by MS.

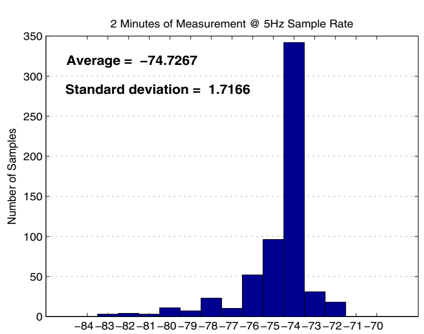

Fig. 8. Short-term measurement from L1.

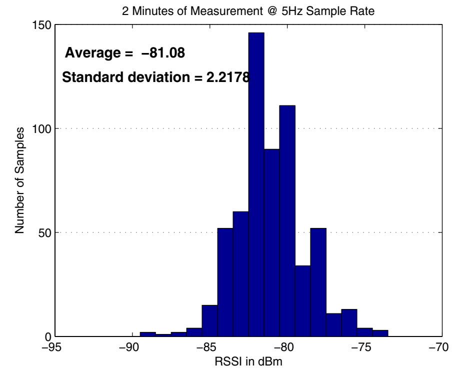

Fig. 9. Short-term measurement from L2.

from statistical viewpoint. Experimentation was designed as collecting RSSI of beacon3 at location L_1 and L_2 respectively. The two locations are separated by 1.2 meters, and nobody resident below L_1 but frequent activities occur around L_2 . The short-term measurement lasts 2 minutes at 5 Hz sample rate and the long-term collection lasts 50 minutes at 0.2 Hz. Fig. 8 and Fig.9 presents the short-term high-rate measurement results. We have concluded below from the histograms:

- L_1 has higher mean value than L_2. The instinctive reason may be L_1 being closer to beacon and signal strength is sensitive to the distance between location and beacon.
- Both of the locations get very small standard deviation in short time. L_2 has larger deviation mainly for people walking in aisle frequently.

The long-term probabilistic distributions are depicted in Fig. 10 and Fig. 11. The two distributions show these characteristics obviously:

- The two distributions have some similarities of mean value and standard deviation, as the signal has property of long-term stability.
- The shape of the long-term distribution is smoother than that of the short-term distribution; this is consistent to the result of [7].
- Left-skewed distributions are prominent with both short-term and long-term measurement that could be approximated by a lognormal distribution; this is consistent to the result of [9].

It is concluded that the statistical characters of signal from one beacon ensure the feasibility and practicability. However, the system's performance depends on the separation of location fingerprints. To investigate how the pattern of fingerprinting at different locations effects the location separation, we carry out the experimentation at location L_1 and L_2 by measuring signal strength from beacon 2 and beacon3 at the same time. 100 groups of data are sampled in 20 seconds. Fig. 12 plots the frequency of occurrence of each sample pattern. The plot shows most of data is surrounding the centers. L_1 has more sharper peak for the its peaceful surroundings. It suggests that only two beacons

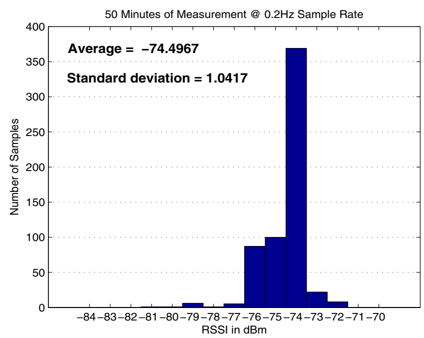

Fig. 10. Long-term measurement from L1.

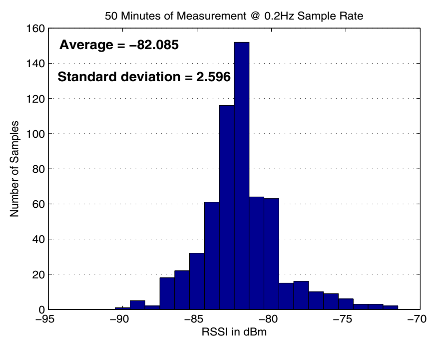

Fig. 11. Long-term measurement from L2.

are sufficient for distinguishing the two locations. The valley between them is very clear that the edge can be used by the algorithm as boundary of classification. We believe adding more beacons can promote the performance greatly. So in system implementation, we add beacon 1 and beacon 4 as Fig. 6 shows.

Fig. 13 shows how the mean RSSI (10 samples @ 5 Hz) varies at 30 calibration points. Compared with Fig. 2 in [5], we find our results fluctuate frequently. It suggests that in office environment, it is more sophisticated than corridor. When the BS is about 6 meters away from the beacon, the mean value fluctuates with ± 5 dBm as dashed shows because of the number of reflect path increases greatly in such points. But in positive points, the RSSI is sensitive to locations in such environment.

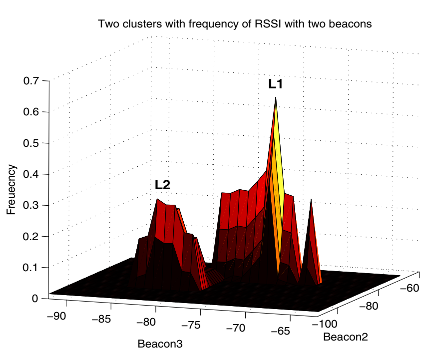

Fig. 12. Separation of location fingerprints.

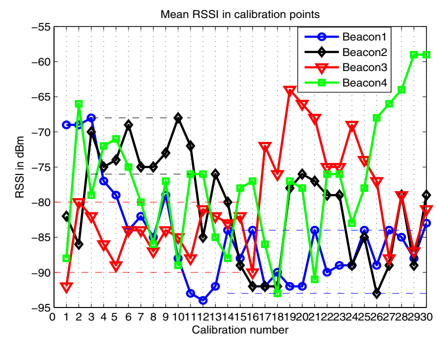

Fig. 13. Mean RSSI in calibration points .

D. Offline Training Phase

The location fingerprinting is collected at each point of the 30 calibration points. The probabilistic distributions of four directions are obtained by (1) where we use the sample rate 5 Hz and 10 samples per point.

E. Online Estimation Phase

We pick points around the calibration points randomly to perform the online estimation phase, which contains two steps for each location. First, measure and average the RSSI from beacons. We take 4 samples per beacon at 5 Hz sample rate to insure this step will be finished in less than 4 seconds. The average RSSI forms the observation tuple $O_{xy} = (o_{xy}^1, o_{xy}^2, o_{xy}^3, o_{xy}^4)^T$ and be applied in (4) (5) (6) to triangulate location L_{ij}. Second, search the fingerprinting database which stores the prior probability $h_{ij}^k(\zeta)$ to find the (i, j) which makes $P(L_{ij}|O_{xy})$ maximized.

F. Results and Analysis

Our experiments returned 70% correct locations average with the tolerance of 0.5 meters. Most false estimation happen when someone was standing near the mobile station or crossing fast under the beacon. Another important phenomenon confused us is that, although the TI's design note indicates omni-directional radiation pattern [23], the direction of the module which connected to the laptop impact the RSSI obviously. We suspect it as the battery or LCD's influence because they are the only obstructs that block the radiation of the antenna. However, this will be analyzed in future.

IV. CONCLUSIONS AND FUTURE WORK

A new location estimation technique based on fingerprinting in ZigBee network has been introduced. Through collecting received signal strength from several beacons to form training dataset and comparing the mobile station's online collected signal with using Bayesian inference, the system can triangulate the location within 70% accuracy with the tolerance of 0.5 meters that is quite encouraging.

Compared with other indoor location systems based on WLAN [5]-[9], [15]-[19], this design has the similar system methodology and even the same algorithm [7]-[9], [17]-[19]. But this work first extends the method into office environment by the easy network construction. And our design is not only fit for indoor but also outdoor environment with much lower cost. It can provide a set of embedded location-aware applications such as wireless sensor network.

Our work can be extended in several directions. Several pattern classification methods can be applied and evaluated, such as k-Nearest Neighbor [5], Neural Network [20], and Support Vector Machine [20]. Direction sensors can be added to reduce search space of the fingerprinting database such as digital compass.

REFERENCES

[1] L. Yang and F.-Y. Wang, "Driving into Intelligent Spaces with Pervasive Communications", *IEEE Intelligent Systems* , vol. 22 , no. 1, pp. 12-15, Jan.-Feb. 2007.

[2] T. P. Moran and P. Dourish, "Introduction to This Special Issue on Context-Aware Computing", *Human-Computer Interaction* , vol. 16, pp. 87-95, 2001.

[3] K. Pahlavan, X. Li, and J. P. Makela, "Indoor Geolocation Science and Technology", *IEEE Communications Magazine* , vol. 40, no. 2, pp. 112-118, Feb. 2002.

[4] J. Hightower and G. Borriello, "Location Systems for Ubiquitous Computing", *IEEE Computer* , vol. 34, no. 8, pp. 57-66, Aug. 2001.

[5] P. Bahl and V. N. Padmanabhan, "RADAR: An In-building RF-based user Location and Tracking System", in *Proc. IEEE Nineteenth Annual Joint Conference Computer and Communications Societies (INFOCOM'00.)* , Tel Aviv, Israel, pp. 775-784, Mar. 2000.

[6] J. Small, A. Smailagic, and P. Siewiorek, Daniel, "Determining User Location for Context Aware Computing Through the Use of a Wireless LAN Infrastructure", *ACM Mobile Networks and Applications* , vol. 6, 2001.

[7] Z. Xiang, S. Song, J. Chen, H. Wang, J. Huang, and X. Gao, "A wireless LAN-based Indoor Positioning Technology", *IBM Journal of Research and Development* , vol. 48, no. 5/6, pp. 617-626, Sept./Nov. 2004.

[8] M. A. Youssef, A. Agrawala, and S. A. Udaya, "WLAN Location Determination via Clustering and Probability Distributions", in *Proc. IEEE International Conference on Pervasive Computing and Communications(PerCom 2003)* , Dallas-Fort Worth, TX, pp. 143- 150, Mar. 2003.

[9] K. Kaemarungsi, "Design of Indoor Positioning Systems Based on Location Fingerprinting Technique", *Ph.D.Thesis* , University of Pittsburgh, 2005.

[10] A. Wheeler, "Commercial Applications of Wireless Sensor Networks Using ZigBee", *IEEE Communications Magazine* , vol. 45, no. 4, pp. 70-77, Apr. 2007.

[11] IEEE Std. 802.15.4, "Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs)", May 2003.

[12] ZigBee Alliance, "ZigBee Specification Ver. 1.0", Dec. 2004.

[13] Jennic Inc. (2007, April) "ZigBee Stack User Guide" [Online]. Available:http://www.jennic.com/support/viewfile.php?fileID=0000000044.

[14] C. Gentile and L. Klein-Berndt, "Robust Location Using System Dynamics and Motion Constraints", in *Proc. IEEE International Conference on Communications* , vol. 3, pp. 1360-1364, Jun. 2004.

[15] T. S. Rappaport, *Wireless Communicaions: Principles and Practice* , 1st ed. Upper Saddle River, NJ: Prentice Hall PTR, 1996.

[16] S. Ali and P. Nobles, "A Novel Indoor Location Sensing Mechanism for IEEE 802.11 b/g Wireless LAN", in *Proc. The Fourth Workshop on Positioning, Navigation and Communication (WPNC '07)* , pp. 9-15, Mar. 2007.

[17] 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* , vol. 9, pp. 155-164, 2002.

[18] A. Kushki, K. N. Plataniotis, A. N. Venetsanopoulos, and C. S. Regazzoni, "Radio Map Fusion for Indoor Positioning in Wireless local Area Networks", in *Proc. The Eighth International Conference on Information Fusion* , vol. 2, pp. 8, Jul. 2005.

[19] A. M. Ladd, K. E. Bekris, G. Marceau, A. Rudys, L. E. Kavraki, and D. S. Wallach, "Robotics-based Location Sensing Using Wireless Ethernet", in *Proc. ACM International Conference on Mobile Computing and Networking (MOBICOM'02)* , pp. 227-238, 2002.

[20] R. Battiti, M. Brunato, and A. Villani. (2002, Oct.) "Statistical Learning Theory for Location Fingerprinting in Wireless LANs", [Online]. Available: http://rtm.science.unitn.it/"battiti/archive/86.pdf.

[21] T. King, S. Kopf, T. Haenselmann, C. Lubberger, and W. Effelsberg, "COMPASS: A Probabilistic Indoor Positioning System Based on 802.11 and Digital Compasses", in *Proc. The First International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization(WiNTECH '06)* , pp. 34-40, Sep. 2006.

[22] Texas Instruments Inc. (2007, April) "Design Note DN007 2.4 GHz Inverted F Antenna".