# Indoor Positioning Techniques Based on Wireless LAN

Binghao Li[1], James Salter[2], Andrew G. Dempster[1] and Chris Rizos[1]
[1]*School of Surveying and Spatial Information Systems, UNSW, Sydney, 2052, Australia*
[2]*School of Computer Science and Engineering, UNSW, Sydney, 2052, Australia*
*Email: binghao.li@student.unsw.edu.au, jsalter@cse.unsw.edu.au*

## Abstract

*As well as delivering high speed internet, Wireless LAN (WLAN) can be used as an effective indoor positioning system. It is competitive in terms of both accuracy and cost compared to similar systems. To date, several signal strength based techniques have been proposed. Researchers at the University of New South Wales (UNSW) have developed several innovative implementations of WLAN positioning systems. This paper describes the techniques used and details the experimental results of the research.*

## 1. Introduction

Location Based Services (LBS) are mobile applications which rely on a user's location to deliver context aware functionality. Industry forecasts for this area predict huge market growth and revenue. One of the key issues for LBS is the positioning technology. GPS is the most popular positioning system; however, it is not suitable for indoor positioning. Purpose built systems such as Active Badge, cricket, The Bat etc., can be used in indoor environments [5]. However, for cost reasons, people prefer to use existing infrastructure such as mobile phone networks [4], television signals [14], and wireless LAN (WLAN). Of these, WLAN can be implemented with the least effort, as its associated consumer hardware is the most readily available. It is also the most accurate, as the signal strength displays high spatial variance, and WLAN chipsets are relatively easily programmed for this purpose.

WLAN aims to provide local wireless access to fixed network architectures. Its market is growing rapidly as the flexibility, connectivity, mobility, and low cost of this technology meets the needs of consumers. A group of specifications has been ratified by IEEE 802.11 working group. Of these, 802.11b (also known as "Wi-Fi") has become the industry standard. It operates up to 11 Mbps in the 2.4 GHz band, which is the only accepted ISM band available worldwide [6].

Obviously, WLAN is not designed and deployed for the purpose of positioning. However, measurements of signal strength (SS) of the signal transmitted by either access point (AP) or station imply the location of any mobile user (MU). Many SS based techniques have been proposed for position estimation in environments in which WLAN is deployed. There are essentially two categories of such techniques. One uses a signal propagation model and information about the geometry of the building to convert SS to a distance measurement. With knowledge of the coordinates of the WLAN APs, the method of trilateration can then be used to compute the position of the mobile user. The other category of WLAN positioning is known as location fingerprinting. The key idea behind fingerprinting is to map location-dependent parameters of measured radio signals in the area of interest. In WLAN the location-dependent parameter is the received signal strength indicator (RSSI) at the APs or MU [3] [16], which can be extracted from the 802.11 chipset through low level APIs.

Increasingly, WLAN positioning systems (WPS) are seen as a convenient position determination technique for indoor environments, or in downtown urban areas, wherever WLAN is deployed. Researchers at the University of New South Wales (UNSW) have developed several innovative implementations of a WPS. This paper describes the techniques used and details the experimental results of the research.

## 2. Trilateration approach

The trilateration approach is simple. Three base stations (or more) with known coordinates are required. If the distance r from the AP to a MU can be measured, a circle with radius r can be drawn. Circles intersect at one point which is the position of MU. However, the measurements obtained are SS rather than the distance. Hence, the SS should be converted to a distance measurement first. So, the trilateration approach consists of two steps: the first step, using a signal propagation model to convert SS to AP-MU separated distance; the second step, least squares or another method (such as the geometric method) can be used to compute the location. The first step is the key of this approach.

Since the environment varies significantly from place to place, the simplest way to find the relationship of SS and the separated distance is collecting some SS data at some points with the known coordinates. This means an extra procedure, named a learning procedure, has to be added to the trilateration approach. In the experiment, data was collected to determine the RF propagation model; precisely the relationship of AP-MU separated distance and received SS (RSS) of AP. The experiment shows the accuracy of this approach is about 4-5 metres (using a general model for all the APs' signal).

To improve the accuracy of the trilateration approach, a hybrid method was proposed [15]. This method is based on the fact that in small localities, such as in a room, the propagation model is better behaved. This method has two stages: in stage one, find the small area the MU is in; in stage two, using trilateration to accurately estimate the location of MU. The experiment shows that the hybrid method can improve the accuracy significantly. However, it is still slightly worse than using fingerprinting with a medium training phase.

The difficulty with the trilateration approach is obtaining the distance measurement from the SS accurately. Indoor radio signal propagation is very complicated, because of signal attenuation due to distance, penetration losses through walls and floors, and the effect of multipath propagation. Interference from other signals is also a problem. 802.11b uses the same frequency band as that used by microwave ovens, cordless phones, and Bluetooth devices etc. Hence, in the 2.4GHz frequency band, those devices can be sources of interference. Furthermore, the orientation of the receiver's antenna, and the location and movement of people inside the building, can affect the SS significantly [9]. It is extremely difficult to build a sufficiently good general model of signal propagation that coincides with the real world situation. Hence the fingerprinting approach is more attractive.

## 3. Fingerprinting approach

Location fingerprinting consists of two phases: 'training' and 'positioning'. The objective of the training phase is to build a fingerprint database. In order to generate the database, reference points (RPs) must first be carefully selected. Locating a MU at one RP location, the SSs of all the APs are measured. From such measurements the characteristic feature of that RP (its SS) is determined, and is then recorded in the database. This process is repeated at another RP, and so forth until all RPs are visited. In the positioning phase, the MU measures the RSS at a place where it requires its position. The measurements are compared with the data in the database using an appropriate search/matching algorithm. The outcome is the likeliest location of the MU. The whole process is illustrated in Figure 1.

The fingerprinting approach has been accepted as an effective method for WLAN positioning although there are still a lot of problems. There are in fact two ways to estimate the unknown location. The simplest one is the deterministic method [3] [10]. The average SS of each WLAN AP measured at each RP is used to create the fingerprint database. Since the variation of the SS measured at each point is large, in order to achieve more accurate results, the probabilistic approach has also been developed [9] [11]. Unfortunately, the distribution of the SS is non-Gaussian. Even worse, it varies at different locations, and at the same location when the orientation of the antenna changes [9] [12]. Hence many measurements are necessary, and this takes more time to generate the RSS distribution at each RP. Furthermore, this increases the database size and the computational burden. Nevertheless, the establishment of the location fingerprint database is an essential prerequisite. To achieve a good estimation of user location, the more RPs, or in other words, the smaller the granularity, the better. And the more measurements obtained at each point the better. However, more RPs and more measurements mean that the training phase is a significant task in terms of labour and time.

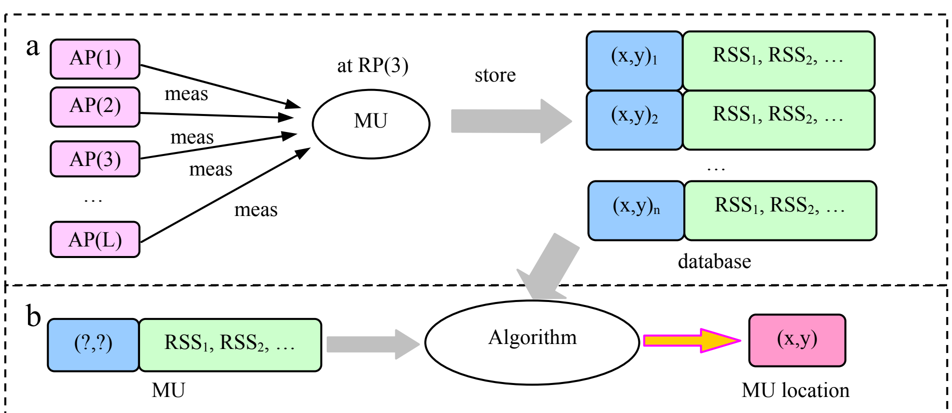


**Figure 1. Two phases of fingerprinting: (a) training phase and (b) positioning phase**

## 4. Deterministic method of Fingerprinting

Using this method, the structure of the fingerprint database is relatively simple and the feature of the RP is only determined by the average RSSs of each AP's (refer to Figure 1 (a)). Many algorithms can be used to estimate the position of MU. The basic one is nearest neighbour (NN).

First, the signal distance between the measured SS vector [s1 s2 … sn] and the SS vector in the database [S1 S2 … Sn] is computed. The generalized distance between two vectors is

$$L_q = \left( \sum_{i=1}^{n} |s_i - S_i|^q \right)^{\frac{1}{q}} \tag{1}$$

Manhattan distance and Euclidean distance are L1 and L2 respectively. Experiments show using Manhattan distance can obtain the best accuracy (although the improvement is not significant) [16]. The nearest neighbour is the point with the shortest signal distance.


**Figure 2. Experimental test bed**

If K (K≥2) nearest neighbours (those with the shortest distance) (KNN) are taken into account, the average of the coordinates of K points can be used as the estimate of the MU location. Similar with KNN, but the weighting scheme is used in KWNN (K weighted nearest neighbour, K≥2). When the location of MU is computed the weighted average is calculated rather than the average. One of the weighing schemes is using the inverse of the signal distance as the weight. Other algorithms such as the smallest polygon [13] and neural networks [7] can also be used.

An experiment was carried out to evaluate the fingerprinting approach. The experimental test bed is shown in Figure 2. Five APs were installed at the locations (pentagram symbols) and there are in total 132 RPs (crosses), and 30 test points (squares).

To investigate the effect of the granularity, the number of RPs was intentionally reduced to 99, 66, 33, and 16. But the RPs were still spread as evenly as possible in the test area. Hence in total 5 fingerprinting databases were generated. Different algorithms were applied to compute the locations of the 30 test points based on a different size of the database. In KNN, K equals 2, 3, 4, 5 or 6. In KWNN, K equals 2, 3, 4 or 5. Table 1 lists all the mean distance errors computed using the different algorithms for the different cases. In general, the KNN and KWNN can achieve better accuracy than the simple NN algorithm. Nevertheless, when the granularity of the RPs is large, the NN even performs better than some of the more complicated algorithms. When KNN is used, in general K equals 3 or 4 will yield the best result. This indicates that only using the two nearest neighbours is not enough (some of the useful information has been ignored), but too many nearest neighbours could decrease the accuracy of the estimator since some of the nearest neighbours are too far from the estimated points. KWNN slightly improves the accuracy of estimation. But none of these algorithms can always provide the best result.

Figure 3 shows the average of the positioning error using all the algorithms listed in Table 1. Apparently, when the average granularity reduces (or the number of RPs increases), the accuracy of the MU's location estimate increases. But when the density of the RPs is high, the rate of increase of the accuracy decreases. Increasing the number of RP does not in itself ensure high accuracy positioning when the granularity is already adequate. But it is very hard to estimate the degree of granularity of RPs required to ensure a specific accuracy since in real world situations the signal propagation environment may vary significantly.

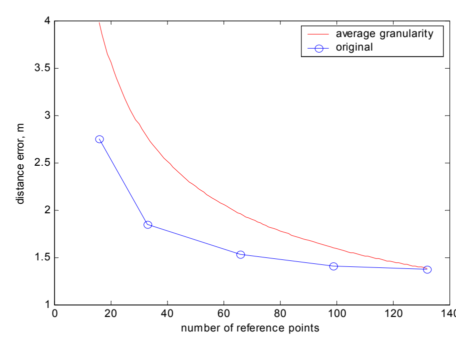


**Figure 3 Mean of average distance errors using different number of RPs**

The demerit of the fingerprinting approach is apparent: the training phase is a significant task (no matter which of the deterministic or probabilistic method are used) if the very accurate positioning results are to be achieved. Furthermore, any changes to the environment that would affect signal strength necessitate re-training. If many RPs and a lot of measurements of each RPs are required, the application of the fingerprinting approach would be limited simply because of the inconvenience. Obviously, effort should be made to overcome this problem.

Table 1. Mean distance error using different algorithm for different cases (unit: m)

| | NN | 2NN | 3NN | 4NN | 5NN | 6NN | 2WNN | 3WNN | 4WNN | 5WNN |
|---|---|---|---|---|---|---|---|---|---|---|
| **Test1 (132 RPs)** | 1.75 | 1.47 | 1.29 | 1.23 | 1.38 | 1.31 | 1.49 | 1.29 | 1.19 | 1.31 |
| **Test2 (99 RPs)** | 1.63 | 1.52 | 1.38 | 1.31 | 1.36 | 1.39 | 1.53 | 1.37 | 1.27 | 1.30 |
| **Test3 (66 RPs)** | 1.74 | 1.47 | 1.51 | 1.60 | 1.52 | 1.60 | 1.48 | 1.44 | 1.49 | 1.43 |
| **Test4 (33 RPs)** | 1.78 | 1.93 | 1.94 | 1.72 | 1.99 | 2.12 | 1.79 | 1.79 | 1.64 | 1.75 |
| **Test5 (16 RPs)** | 2.55 | 2.34 | 2.65 | 2.98 | 3.41 | 3.99 | 2.11 | 2.28 | 2.45 | 2.69 |

The conventional method of generating the database does not utilise the spatial correlation of measurements sampled at adjacent RPs. When measurements at a small number of RPs are made, they not only provide information at these points, but also imply information of the surrounding area. If a denser database can be generated efficiently by interpolation based on a small number of RPs, labour effort and time can be saved during the training phase [16]. Two methods, inverse distance weighting (IDW) and universal kriging (UK), are chosen to generate the database. The methodology is illustrated in Figure 4.

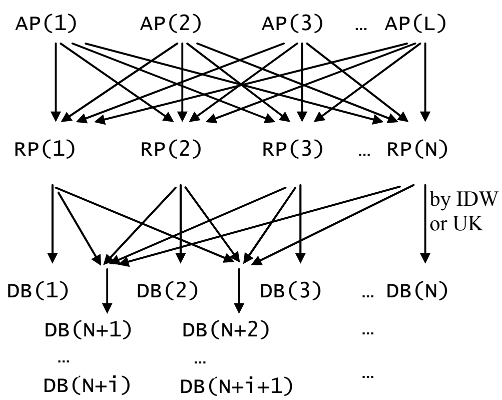


Figure 4 Proposed database generation methodology

The data collected in the experiment for the fingerprinting approach was further utilized to evaluate the proposed method. Three groups of databases were generated. They are the original group (using the conventional method), the IDW group (database based on IDW), and the kriging group (database based on UK). In each group, there are five different versions of the databases.

Slightly differently from the test of fingerprinting approach, five algorithms have been selected here: the NN, KNN (K=3,4) and KWNN (K=3,4). Hence five distance errors would be generated when MU's location was estimated. The mean of the five distance errors are computed to compare the different methods. Figure 5 shows the positioning error using the different databases: the original databases, the databases generated by the IDW and the ones generated by UK.

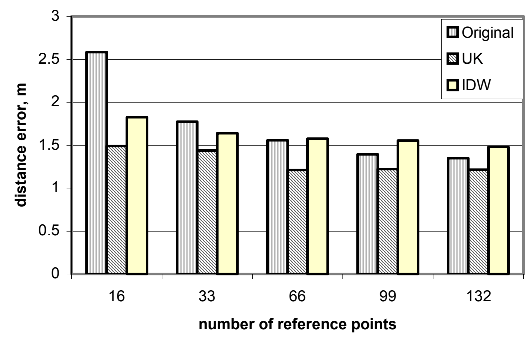


Figure 5 Mean of average distance errors using different databases

From Figure 5 two conclusions can be drawn. Firstly, UK can efficiently estimate the RSS using the information of some of the RPs. This means that UK can yield a database of location fingerprints with good quality. Secondly, when the density of RPs reaches a particular value (in this case around 66 points), kriging cannot provide better estimation. On the other hand, it is unnecessary to measure so many RPs (66 here) to achieve the best estimation. This implies the training phase can be shortened significantly. In the worst case, when only very few RPs can be measured, kriging can also obtain reasonable location estimates. In this experiment, when only 16 RPs were used, the estimation error is less than using the original database of 66 RPs, and only slightly worse than 99 or 132 RPs. One thing must be emphasised is that 16 RPs means in small room (of around 20m2) there is only one RP, and in a large room (of around 50m2) only 2 or 3 RPs are required. This makes the training phase very flexible, and if the environment is changing, fast training can be carried out and a new database of location fingerprints can be generated quickly.

## 5. Probabilistic method of Fingerprinting

While the deterministic method gives reasonable localization accuracy, it discards much of the information present in the training data. Each fingerprint summarizes the data as the average signal strength to visible access points, based on a sequence of signal strength values recorded at that location. However, signal strength at a position can be characterized by more parameters than just the average.

Figure 6 shows the structure of the signal strength for a single fingerprint recording. Both the average signal strength and the variance differs for each access point. The signal strength distribution is also not necessarily Gaussian. We would like to consider as much of this information as possible when performing comparisons between the input and the signal strength map, to maximize accuracy.

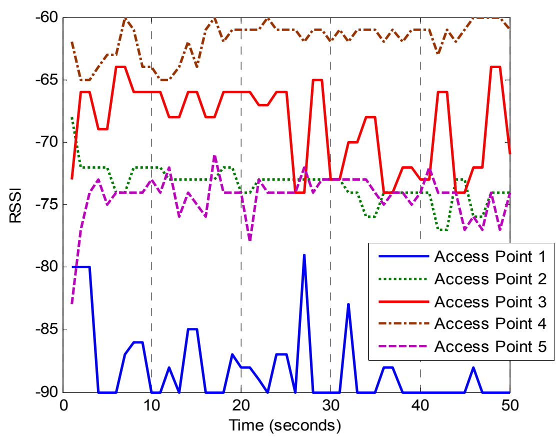


**Figure 6 Signal strength recorded at a single location and orientation**

This led researchers to consider a Bayesian approach to WLAN localization [8] [9]. This had been employed with some success in the field of robot localization [1] [2].

For localization, Bayes rule can be written as

$$p(l_t|o_t) = p(o_t|l_t)p(l_t) \cdot N \tag{2}$$

where lt is a location at time t, ot is an observation made at t (the instantaneous signal strength values), and N is a normalizing factor that ensures that all probabilities sum to 1. In other words: the probability of being at location l given observation o is equal to the probability of observing o at location l, and being at location l in the first place. During localization, this conditional probability of being at location l is calculated for all fingerprints. The most likely location is then the localizers output.

To calculate this, it is necessary to calculate the two probabilities on the right hand side of the equation. p(ot|lt) is known in Bayesian terms as the likelihood function. This can be calculated using the signal strength map. For each fingerprint, the frequency of each signal strength value is used to generate a probability distribution as the likelihood function. The raw distribution can be used, but as it is typically noisy and incomplete, the data is usually summarized as either a histogram, with an empirically determined optimal number of bins, or as a discretized Gaussian distribution parameterized using mean and standard deviation. Other representations are also possible; the Bayesian approach allows us to use any algorithm capable of generating a probability distribution across all locations.

In its simplest incarnation, the Bayesian localizer calculates the prior probability p(lt) as the uniform distribution over all locations. This encodes the idea that, before each attempt at localization, the target is equally likely to be at any of the locations in the fingerprint map. In order to achieve higher accuracy, we can calculate this probability using our knowledge of the target's likely motion, historical information from previous user habits, collision detection, and anything else that affects the prior probability that can be modelled probabilistically. Markov Localization [1] suggests using the transitional probability between locations. This probability is described as

$$p(l_t) = \sum_{l_{t-1}} p(l_t|l_{t-1})p(l_{t-1}) \tag{3}$$

In other words, p(l) is the sum of the transitional probability from all locations at t-1 to l at the current time t, multiplied by the probability of being at those locations at t-1. $p(l_{t-1})$ is known from previous localization attempts. We calculate $p(l_t|l_{t-1})$ using a motion model, the details of which are specific to how we expect the target to move. For a walking person, the simplest and most effective approach is to calculate the probability based on how far the user can move between t and t-1. This is not very far if our localizer runs once every second. In practice, this extra calculation serves to remove noise affected outliers from the output.

In practice, the Bayesian localizer proves more accurate than the nearest neighbour technique because it takes into account more information from the training data and filters the output using a motion model. Figures 7 and 8 show the performance of the Bayesian technique for static and mobile localization, respectively. Static localization is performed for targets not expected to move, and takes the prior probability as the uniform distribution. For this test, the Bayesian localizer marginally outperformed the nearest neighbour technique. For mobile localization, the prior probability was calculated using a simple motion model, which caused the accuracy to be significantly improved compared to the nearest neighbour approach. The median error when summarizing signal strength information as a Gaussian is approximately 1.4 metres, compared to 2.2 metres for the NN in this test. As well as an improved median error, it is clear that for both static and mobile localization, the

90th percentile error is significantly reduced when using the Bayesian approach (3.5 metres against 5.1 metres for mobile localization). This suggests that the Bayesian approach is more reliable.

As well as improving accuracy, the Bayesian localizer provides a framework for the integration of other sensors present. Any sensor for which p(o|l) can be calculated, such as infra-red or mobile phone signal strength, can be integrated into the model by running the same Bayesian update equation on a shared probability distribution. Relative sensors such as wheel encoders and inertial sensors can be included by using them to inform the motion model used to calculate the prior probability, as they can be used to calculate the transitional probability (how far a distance has been travelled) with a high degree of accuracy.

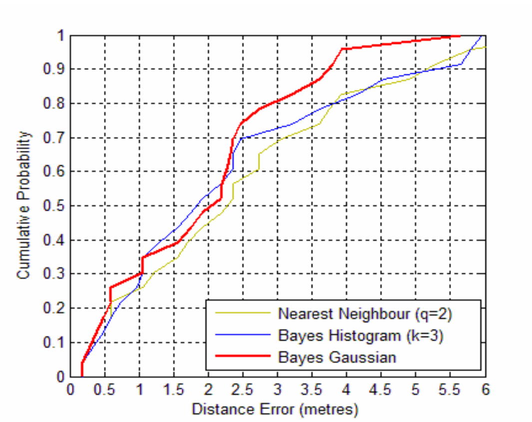


**Figure 7 Cumulative Error probability for static localization**

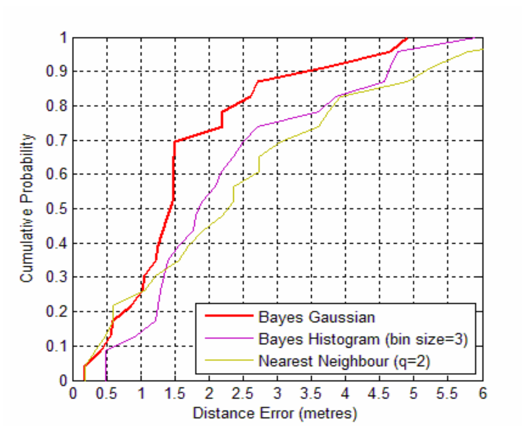


**Figure 8 Cumulative Error probability for mobile localization**

## 6. Concluding Remarks

In this paper, the techniques used for WLAN indoor positioning have been discussed. The advantages and drawbacks of each technique have been compared based on the experiment results. Generally, using fingerprinting can achieve few meters accuracy. To overcome the drawbacks of this technique, more efforts are needed especially for the probabilistic method in the future research.

## 7. References


[1] Simmons, R., & Koenig, S., "Probabilistic robot navigation in partially observable environments", *The International Joint Conference on Artificial Intelligence (IJCAI'95)*, Montreal, Canada, 20-25 August 1995, pp. 1080-1087.

[2] Fox, D., *Markov Localization: A probabilistic Framework for Mobile Robot Localization and Navigation*, PhD thesis, Institute of Computer Science III, University of Bonn, Germany, 1998.

[3] Bahl, P., & Padmanabhan, V.N., "RADAR: An in-building RF-based user location and tracking system", *IEEE Infocom 2000*, Tel Aviv, Israel, 26-30 March 2000, vol. 2, pp. 775-784.

[4] Laitinen, H., (ed), *Cellular Location Technology*, CELLO project report, 2001. http://www.vtt.fi/tte/tte35/pdfs/CELLO-WP2-VTT-D03-007-Int.pdf

[5] Hightower, J., & Borriello, G., "Location systems for ubiquitous computing", *IEEE Computer*, August 2001, 34, (8), pp. 57-66

[6] Bing, B. *Wireless Local Area Networks: the New Wireless Revolution*, Wiley-Interscience, New York, Chichester, 2002.

[7] Battiti, R., Nhat, T.L., & Villani, A., "Location-aware computing: a neural network model for determining location in wireless LANs", Technical Report DIT-02-0083, February 2002 http://eprints.biblio.unitn.it/archive/00000233/01/83.pdf

[8] Roos, T., Myllymaki, P., Tirri, H., Misikangas, P., & Sievanan, J., "A probabilistic approach to WLAN user location estimation", *International Journal of Wireless Information Networks*, 9(3), July 2002.

[9] Ladd, A.M., Bekris, K.E., Rudys, A., Marceau, G., Kavraki L.E., & Dan, S., "Robotics-based location sensing using wireless Ethernet", *Eighth ACM Int. Conf. on Mobile Computing & Networking (MOBICOM)*, Atlanta, Georgia, US, 23-28 September 2002, pp. 227-238.

[10] Saha, S., Chaudhuri, K., Sanghi, D., & Bhagwat, P., "Location determination of a mobile device using IEEE 802.11b access point signals", *IEEE Wireless Communications & Networking Conference (WCNC)*, New Orleans, Louisiana, US, 16-20 March 2003, vol. 3, pp.1987-1992.

[11] Youssef, M., Agrawala A., & Shankar, A.U., "WLAN location determination via clustering and probability distributions", *IEEE Int. Conf. on Pervasive Computing & Communications (PerCom) 2003*, Fort Worth, Texas, US, 23-26 March 2003, pp.143-150.

[12] Wang, Y., Jia, X., Lee, H.K., and Li, G.Y., "An indoor wireless positioning system based on WLAN infrastructure", *6th Int. Symp. on Satellite Navigation Technology Including Mobile Positioning & Location Services*, Melbourne, Australia, 22-25 July 2003, CD-ROM proc., paper 54.

[13] Pandya, D., Jain, R., & Lupu, E., "Indoor location estimation using multiple wireless technologies", *14th IEEE Int. Symp. on Personal, Indoor, & Mobile Radio Communications (PIMRC)*, Beijing, China, 7-10 September 2003, vol. 3, pp. 2208-2212.

[14] Rabinowitz, M., & Spilker, J.J., Jr., "A new positioning system using television synchronization signals", *Rosum Corporation white paper*, 2004. http://www.rosum.com/rosum_whitepapers.html

[15] Li, B., Dempster, A.G., Rizos, C., & Barnes, J., "Hybrid method for localization using WLAN", *Spatial Sciences Conference*, Melbourne, Australia, 12-16 September 2005, 341-350, CD-ROM procs.

[16] Li, B., Wang, Y., Lee, H.K., Dempster, A.G. & Rizos, C. "Method for Yielding a Database of Location Fingerprints in WLAN", Communications, IEE Proceedings, Vol. 152, Issue 5, October 2005, pp. 580-586.Li, B., Wang, Y., Lee, H.K., Dempster, A.G. & Rizos, C. "Method for Yielding a Database of Location Fingerprints in WLAN", *Communications, IEE Proceedings*, Vol. 152, Issue 5, October 2005, pp. 580-586.