Indoor Positioning System Using Artificial Neural Network

Hamid Mehmood, Nitin K. Tripathi and Taravudh Tipdecho
Department of Remote Sensing and Geographic Information Systems,
School of Engineering and Technology, Asian Institute of Technology,
P.O. Box 4, Klong Luang, Pathumthani 12120, Thailand

Abstract: Problem statement: Location knowledge in indoor environment using Indoor Positioning Systems (IPS) has become very useful and popular in recent years. A number of Location Based Services (LBS) have been developed, which are based on IPS, these LBS include asset tracking, inventory management and security based applications. Many next-generation LBS applications such as social networking, local search, advertising and geo-tagging are expected to be used in urban and indoor environments where GNSS either underperforms in terms of fix times or accuracy, or fails altogether. To develop an IPS based on Wi-Fi Received Signal Strength (RSS) using Artificial Neural Networks (ANN), which should use already available Wi-Fi infrastructure in a heterogeneous environment. **Approach:** This study discussed the use of ANN for IPS using RSS in an indoor wireless facility which has varying human activity, material of walls and type of Wireless Access Points (WAP), hence simulating a heterogeneous environment. The proposed system used backpropagation method with 4 input neurons, 2 output neurons and 4 hidden layers. The model was trained with three different types of training data. The accuracy assessment for each training data was performed by computing the distance error and average distance error. **Results:** The results of the experiments showed that using ANN with the proposed method of collecting training data, maximum accuracy of 0.7 m can be achieved, with 30% of the distance error less than 1 m and 60% of the distance error within the range of 1-2 m. Whereas maximum accuracy of 1.01 can be achieved with the commonly used method of collecting training data. The proposed model also showed 67% more accuracy as compared to a probabilistic model. **Conclusion:** The results indicated that ANN based IPS can provide accuracy and precision which is quite adequate for the development of indoor LBS while using the already available Wi-Fi infrastructure, also the proposed method for collecting the training data can help in addressing the noise and interference, which are one of the major factors affecting the accuracy of IPS.

Key words: Artificial neural networks, indoor positioning systems, Wi-Fi, backpropogation method

INTRODUCTION

Wi-Fi Received Signal Strength (RSS) has drawn great attention in recent years for the development of Indoor Positioning Systems (IPS), because of its distinct advantages like already existing infrastructure and wide coverage area. With numerous new Wi-Fi nodes being added daily, the global Wi-Fi coverage is continuously growing. Recent tests have shown that indoor positioning with Wi-Fi RSS based Positioning Systems (WPS) can achieve positioning accuracy of 1 to 4 meter indoors and 10-40 m in the outdoor environment (Prasithsangaree *et al*., 2002; Doherty *et al*., 2001; Hightower and Borriello, 2001).

A typical WPS consists of a Mobile User (MU), wireless Access Point (AP) and a positioning algorithm. There are three basic methods to compute the location of a MU in indoor envrionment; (a) Angle of Arrival (AoA), (b) Time of Arrival (ToA)/Time Difference of Arrival (TDoA) (Yamasaki *et al*., 2005) and (c) Received Signal Strength (RSS) (Kolodziej and Hjelm, 2006).

AoA is based on the time of flight of electromagnetical waves, which requires very accurate timming measurements and thus implementation of AoA requires additional hardware. ToA and TDoA are based on the measurement of the time flight of the signal, the implementation of this technique requires

Corresponding Author: Hamid Mehmood, Department of Remote Sensing and Geographic Information Systems,
School of Engineering and Technology, Asian Institute of Technology, P.O. Box 4, Klong Luang,
Pathumthani 12120, Thailand

clock synchronization between the Wireless Access Points (WAP) (Kawauchi *et al*., 2009), where as for RSS pattern matching algorithms are used to compute the location using RSS (Rooney *et al*., 2007; Papapostolou and Chaouchi, 2009a).

The pattern matching algorithms can be classified as deterministic and probabilistic. The deterministic algorithms are based on non-linear methods e.g., artificial neural networks and genetic algorithms, whereas the probabilistic algorithms are based on statistical theories.

Using RSS, fingerprinting is one of the most frequently used techniques for indoor positioning; which consists of two phases. First phase is called the offline phase in which, the RSS from all the AP is recorded at number of location in the building i.e., generating a location matrix with the "fingerprint" of all the AP at each specific location. The next phase is the online phase, where pattern matching algorithms are used to solve the location matrix and calculate the location of MU by comparing the RSS being observed by the MU with the one recorded in the location matrix. Till now mostly probabilistic algorithms have been used as pattern matching algorithm (Kupper, 2005).

Backpropagation learning algorithm, which is an Artificial Neural Network (ANN) based algorithm, is utilized in this study for indoor positioning with the aim to estimate the location of the MU with minimal error. The reason for selecting ANN is because of its robustness against noise and interference which are one of the major factors affecting the accuracy of IPS. The study is organized in three parts, the first part of the study discusses the experimental site, the second part discusses the proposed ANN model and the third part comprises of the results and discussion.

System description and experimental setup: Indoor radio signal propagation is difficult to predict because of number factors, which cause the fading of the radio signal (Kupper, 2005). This fading causes the radio signal to vary. As this study focuses on development of ANN based IPS, which can be later on integrated with GNSS to provide a ubiquitous positioning system (Mehmood and Tripathi, 2009), hence it is necessary to test the ANN in heterogeneous environment. In this study heterogeneous environment is defined as an environment, which shows fading of the signal because of the below given factors:

- Different hardware manufacturers of installed Wireless Access Point (WAP)
- Age of WAP
- Electrical interference

- Construction material of the wall
- Human activity

The experimental area is the Remote Sensing and GIS laboratory located on the second floor of the North Side academic building at Asian Institute of Technology, Thailand. The experimental area consists of four randomly placed WAP, each with a different life time and hardware manufacturer. The details of the WAP are given in Table 1.

The experimental area has 17 desktop computers, with frequent human movement as the laboratory is 24 hr. accessible to the students. One wall of the study is made of concrete, the other two of wood and the forth one made of glass and wood hence providing a suitable heterogeneous environment to perform the experiment. The study area has the dimensions of 8×9 m as shown in Fig. 2.

The study area is mapped in Cartesian coordinates system, with starting coordinate as 0, 0.

Dell Vostro 1320 laptop equipped with Dell Wireless 1397 WLAN Mini-Card (802.11 b/g) is used for scanning and positioning. Trimble handheld distance measurement unit (LRU) with accuracy of ± 2 mm is used to measure the distance and HTC Diamond mobile p3700 equipped with accelerometer is used to keep the apparatus at level. The apparatus used is shown in Fig. 1.

Table 1: Details of WAPs in experimental area

Name	Mac address	Company	Life in years
RS_Lab	00:16:b6:e1:79:cb	LinkSys	2+
Zyxel_1	00:23:f8:ac:92:8b	Zyxel	<1
Zyxel_2	00:02:cf:c7:e1:9c	Zyxel	<1
AIT	00:02:2d:0e:73:b2	LinkSys	5+

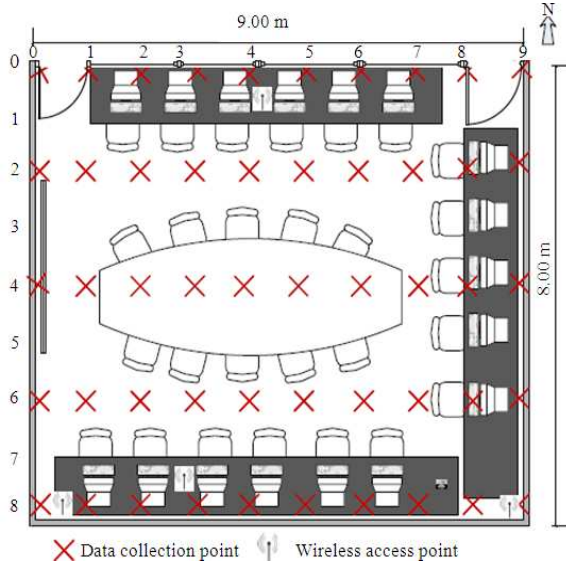Fig. 1: Hardware used for the development of the system

Fig. 2: Experimental site for IPS

Application software architecture: The ANN application is developed using NeuronDotNet, which is an open source engine for the development of ANN based applications. The engine is written in C# and is compatible with the NET platform. The scanning module is developed by modifying inSIDDer, which is also an open source Wi-Fi scanner. The programming language used for development is C#.

MATERIALS AND METHODS

The system consists of two modules:

* Scanning module
* Positioning module

Figure 3 shows the flow of both the modules.

Scanning module: Scanning module or the offline phase of the system generates the training data for the ANN. Scanning module consists of a Wi-Fi RSS scanner and logger.

Using the scanner, 50 scans are performed, shown as data collection points in Fig. 2. Each scan generates a list of mac address with the corresponding RSS for every visible WAP. In the list, for every WAP there is a single mac address entry and numerous RSS entries. Using this list, Seeding Data (SD) or location matrix is generated.
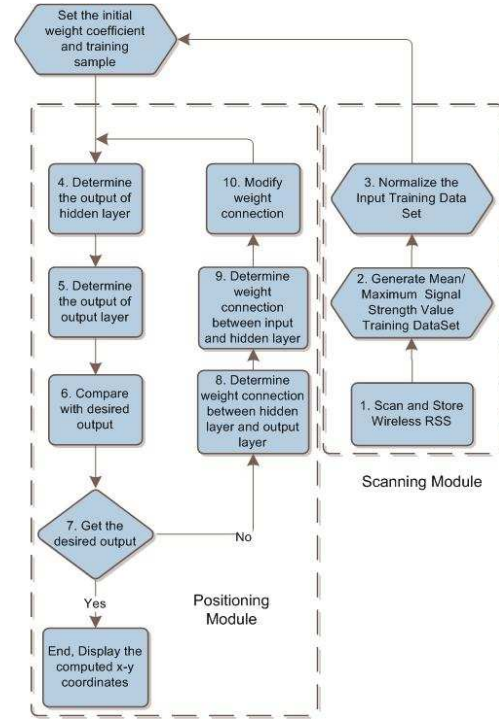

Fig. 3: Flow of the system modules

The SD consists of a unique identifier (mac address of the WAP) and the set of observed RSS corresponding to the mac address. For a scan performed at location k, with i number of visible WAP, the corresponding SD is given by:

$$SD_k = \begin{Bmatrix} (mac_1(ss_1,ss_2,ss_3,...,ss_n)) \\ (mac_2(ss_1,ss_2,ss_3,...,ss_n)) \\ (mac_3(ss_1,ss_2,ss_3,...,ss_n)) \\ . \\ . \\ . \\ (mac_i(ss_1,ss_2,ss_3,...,ss_n)) \end{Bmatrix} \quad (1)$$

Using this SD, two sets of SD are generated i.e., max value SD and mean value SD.

For a scan performed at location k, the max value SD is given by:

$$SD_{max} = \begin{Bmatrix} (mac_1(ss_{max})) \\ (mac_2(ss_{max})) \\ (mac_3(ss_{max})) \\ . \\ . \\ . \\ (mac_i(ss_{max})) \end{Bmatrix} \quad (2)$$

where, ss_{max} is the maximum RSS from the list of observed RSS.

Similarly the mean value SD is given by:

$$SD_{mean} = \begin{Bmatrix} (mac_1(ss_{mean})) \\ (mac_2(ss_{mean})) \\ (mac_3(ss_{mean})) \\ . \\ . \\ . \\ (mac_i(ss_{mean})) \end{Bmatrix} \qquad (3)$$

where, ss_{mean} is the mean of the observed RSS.

During the scanning phase, it is found that the RSS shows variability to change in the direction in which, the scanning is performed. A data sample was collected at coordinate 2, 7 in four directions (North, East, West and South) with only considering Zyxel_2 WAP. The results are shown in the Fig. 4, which show that a relationship exists between the direction in which the scanning is performed and the measured RSS. Data samples taken at different locations indicated the same trend, showing a variation of 5-10 (-dBm). Due to this variation in RSS another SD was generated called directional mean SD. Directional mean SD is generated by taking the mean of RSS recorded in four directions at each location:

$$SD_k = \begin{Bmatrix} (mac_1((d_1)ss_1,ss_2,...,ss_n)) \\ (mac_1((d_2)ss_1,ss_2,...,ss_n)) \\ (mac_1((d_3)ss_1,ss_2,...,ss_n)) \\ (mac_1((d_4)ss_1,ss_2,...,ss_n)) \\ (mac_2((d_1)ss_1,ss_2,...,ss_n)) \\ (mac_2((d_2)ss_1,ss_2,...,ss_n)) \\ (mac_2((d_3)ss_1,ss_2,...,ss_n)) \\ \vdots \\ (mac_i((d_4)ss_1,ss_2,...,ss_n)) \end{Bmatrix} \qquad (4)$$

where, d is the direction in which, the scan is performed. The directional mean SD is given by:

$$SD_{direc\_mean} = \begin{Bmatrix} (mac_1(ss_{direc\_mean})) \\ (mac_2(ss_{direc\_mean})) \\ (mac_3(ss_{direc\_mean})) \\ . \\ . \\ . \\ (mac_i(ss_{direc\_mean})) \end{Bmatrix} \qquad (5)$$

where, ss_{direct_mean} is the mean of observed RSS in four directions.

These three different SDs are used to generate three different training datasets for the ANN. The training dataset consists of the normalized RSS for each WAP and the coordinates at which, the scan was performed. Table 2 shows a sub set of the training dataset generated using directional mean SD.

Positioning module: Neural networks offer different architectures to solve a certain task. Selection of the architecture depends on the kind of the problem to be solved. A good performance is achieved by ANN only when the right number of input layers and hidden layers are selected.

For this study backpropagation method is used to solve the indoor positioning problem. Backpropagation method is a supervised learning method (Fig. 5), where errors propagate backwards from the output nodes to the inner nodes.

The normalized RSS is the input for the ANN whereas, the x and y coordinates is the output. From the scanning phase, 40 scans are used for training phase and 10 scans are used to test the system.
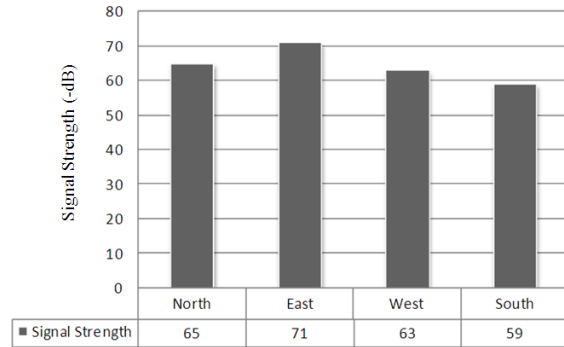

	North	East	West	South
■ Signal Strength	65	71	63	59

Fig. 4: RSS distribution with respect to direction

Table 2: Sub set of training dataset with normalized RSS

AIT	Zyxel_2	RS_Lab	Zyxel1	x	y
0.59	1	0.52	0.5	0.8	0
0.59	0.97	0.53	0.52	0.8	0.1
0.59	0.95	0.5	0.55	0.8	0.2
0.57	0.85	0.5	0.56	0.8	0.3
0.57	0.74	0.5	0.58	0.8	0.4
0.58	0.65	0.51	0.61	0.8	0.5
......	......	......	......	...	...
......	......	......	......	...	...
0.55	0.53	0.52	0.87	0	0.6
0.55	0.57	0.52	0.92	0	0.7
0.55	0.59	0.52	0.95	0	0.8
0.52	0.62	0.59	0.98	0	0.9
0.55	0.68	0.54	1	0	1

The number of hidden layers is calculated using Eq. 6:

$$\left(\text{Number of inputs} + \text{Number of outputs}\right) \times \left(\frac{2}{3}\right) \qquad (6)$$

The properties of the backpropagation method, which give the least error, are given in Table 3.

The error behavior for the directional mean SD, mean value SD and max value SD at 5000 training cycles is shown in Fig. 6 respectively, which shows the directional mean SD reaches the minimum sum squared value at before mean value SD and max value SD.
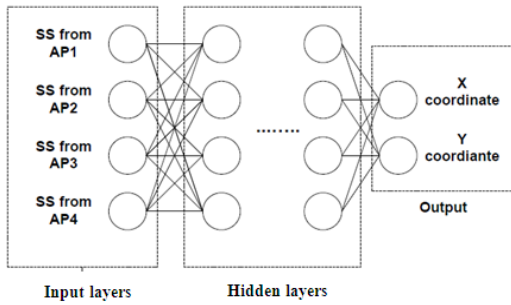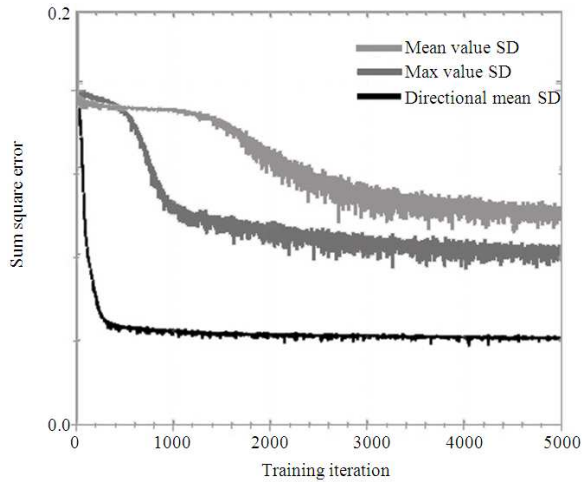

Fig. 5: Backpropogation model

Fig. 6: Error behavior for the directional mean SD, mean value SD and max value SD

Table 3: Properties in ANN model

Property	Value
Input neurons	4
Output neurons	2
Activation function	Sigmoid
Training iterations	5000
Hidden layers	4
Learning rate	0.25

RESULTS AND DISCUSSION

Accuracy and precision are two main parameters which are used to measure the performance of the IPS. Distance Error (DE) is used to measure the precision and vectors maps are used to visualize the precision, whereas Average Distance Error (ADE) is used to measure accuracy.

Precision is defined as the success probability of position estimations with respect to predefined accuracy. The DE is defined as the spatial distance between the calculated position and the real position. DE is calculated using the formula:

$$DE = \sqrt{(x - x_0)^2 + (y - y_0)^2} \qquad (7)$$

The ADE has been calculated using equation:

$$ADE = \frac{1}{n} \sum_{i=1}^{n} DE_i \qquad (8)$$

ADE is the defined as the average of all the summations of the calculated DE. Table 4 shows a comparison of ADE, minimum DE and maximum DE between the position computed using training sets based on directional mean SD, maximum value SD and minimum value SD. Position computed using the directional mean SD show the lowest ADE, which shows that it is more accurate as compared to position calculated using other two training datasets.

Figure 7-9 shows the vector graphs plotted between the observed x-y coordinates and the computed x-y coordinates using training sets based on directional mean SD, maximum value SD and minimum value SD.
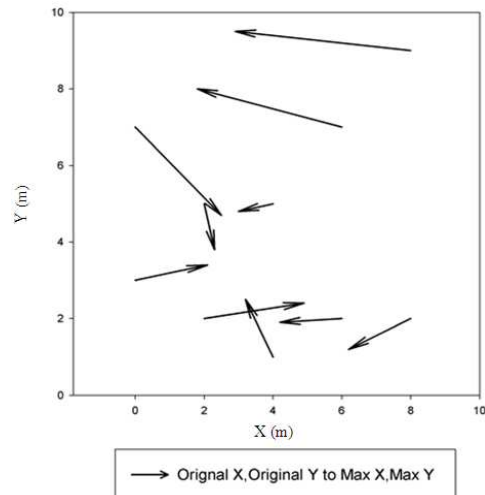

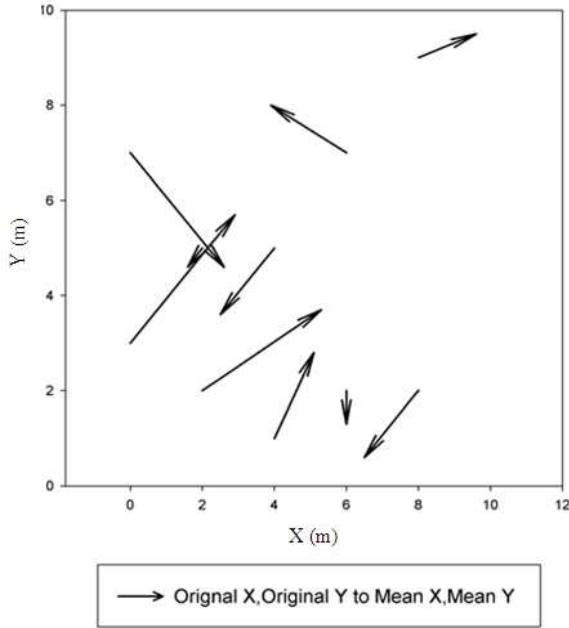Fig. 7: Vector graph for max value training data

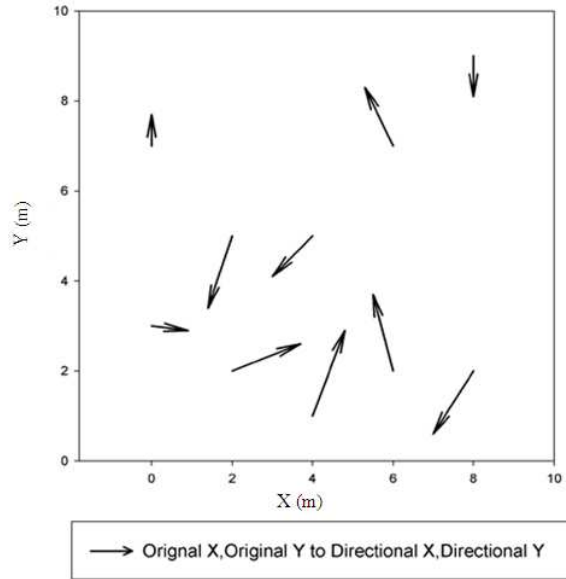Fig. 8: Vector graph for mean value training data

Fig. 9: Vector graph for directional mean value training data

The starting point of the vector is the original x-y coordinate, where as the ending point is the computed x-y coordinate. Figure 9 shows that the vectors generated using the directional mean training data, which have the least magnitude as compared to vectors generated using max value training data and mean value training data, hence showing that the x-y coordinates computed using the directional mean SD are more precise.

Table 4: Comparison of different training data outputs

Training data	Average distance error (m)	Maximum distance error (m)	Minimum distance error (m)
Directional mean	1.43	2.06	0.70
Mean	2.26	3.96	0.56
Max	2.56	5.12	1.01

Table 5: Error comparison between back propagation and probabilistic model

Model	Average distance error (m)	Maximum distance error (m)	Minimum distance error (m)
Backpropagation model	1.43	2.06	0.7
Probabilistic model	2.54	4.37	1.04

The DE distributions for distance mean SD shows that 30% of the values have the distance error less than 1 m. Only 10% of the values have the distance error more than 2 m. The rest of the 60% values show a distance error between 1-2 m.

The results from the ANN method are also compared with a probabilistic method (Papapostolou and Chaouchi, 2009b). Mathematically the probability model for calculating the indoor position can be represented as:

$$P(l_i \mid O) = \frac{P(O \mid l_i)P(l_i)}{P(O)} \tag{9}$$

Where:
P(l_i|O) = The conditional probability of obtaining observation O at position l_i
P(l_i) = The prior probability of position l_i being the correct position, which is obtained by the directional mean training data
P(O) = The normalizing factor (Xiang *et al*., 2004)

Table 5 shows the ADE, maximum DE and minimum DE comparison between the backpropagation method and probabilistic model while using directional mean training data, which shows that the backpropagation method with directional mean SD is more accurate as compared to a probabilistic model using directional mean SD.

CONCLUSION

This research study presented a design and results for ANN based method for indoor positioning using RSSI. The method was tested with three different kinds of training data set in a heterogeneous environment. About 40 training data samples were used in an 8×9 m sample space. The study concluded that the location of the mobile user could be located in the indoor environment with different levels of accuracy and

precision depending on the kind of training data set used. The x-y coordinates computed using the directional mean SD gave the best accuracy and precision; because it was able to integrate into the seeding data the factors that affect the signal strength in a heterogeneous environment.

The backpropagation method was also compared to a probabilistic model for indoor positioning and performance comparison was made in term of accuracy. It was found that, the back propagation model using directional mean data is more accurate as compared to the probabilistic model using the directional mean data. The ANN can achieve 30% error within 1 m. and 60% error within 1-2 m.

Future work consists of optimizing the ANN by using genetic algorithms for weight computation and testing the algorithm over a larger area to test the robustness of the algorithm.

REFERENCES

Doherty, L., K.S.J. Pister and L. El Ghaoui, 2001. Convex position estimation in wireless sensor networks. Proceeding of the 21th Annual Joint Conference on IEEE Computer Communications, Apr. 22-26, IEEE Xplore Press, Anchorage, AK., USA., pp: 1655-1663. DOI: 10.1109/INFCOM.2001.916662

Hightower, J. and G. Borriello, 2001. Location systems for ubiquitous computing. Computer, 34: 57-66. DOI: 10.1109/2.940014

Kawauchi, K., T. Miyaki and J. Rekimoto, 2009. Directional beaconing: A robust WiFi positioning method using angle-of-emission information. Lecture Notes Comput. Sci., 5561: 103-119, DOI: 10.1007/978-3-642-01721-6_7

Kolodziej, K.W. and J. Hjelm, 2006. Local Positioning Systems LBS Applications and Services. 1st Edn., CRC Press, USA., ISBN: 10: 0849333490, pp: 488.

Kupper, A., 2005. Location-Based Services Fundamentals and Operation. 1st Edn., Wiley, USA., ISBN: 13: 978-0470092316, pp: 286.

Mehmood, H. and N.K. Tripathi, 2009. Real Time Self Mapping Hybrid Positioning System. In: Geospatial Crossroads @ GI_Forum'09, Car, A., G. Griesebner and J. Strobl (Eds.). Wichmann, Salzburg, Austria, ISBN: 13: 978-3879074815, pp: 130-139.

Papapostolou, A. and H. Chaouchi, 2009a. Orientation-based radio map extensions for improving positioning system accuracy. Proceeding of the 2009 International Conference on Wireless and Mobile Computing: Connecting the World Wirelessly, June 21-24, ACM Press, Leipzig, Germany, pp: 947-951. DOI: 10.1145/1582379.1582586

Papapostolou, A. and H. Chaouchi, 2009b. WIFE: Wireless indoor positioning based on fingerprint evaluation. Lecture Notes Comput. Sci., 5550: 234-237. DOI: 10.1007/978-3-642-01399-7_19

Prasithsangaree, P., P. Krishnamurthy and P. Chrysanthis, 2002. On indoor position location with wireless LANs. Proceeding of the 13th IEEE International Symposium on Personal Indoor and Mobile Radio Communications, Sept. 15-18, IEEE Xplore Press, USA., pp: 720-724. DOI: 10.1109/PIMRC.2002.1047316

Rooney, S., K. Gardiner and J.D. Carswell, 2007. Wireless positioning techniques-a developers update. Lecture Notes Comput. Sci., 4857: 162-147. DOI: 10.1007/978-3-540-76925-5_12

Xiang, Z., S. Song, J. Chen, H. Wang and J. Huang *et al.*, 2004. A wireless LAN-based indoor positioning technology. IBM J. Res. Dev., 48: 617-626. DOI: 10.1147/rd.485.0617

Yamasaki, R., A. Ogino, T. Tamaki, T. Uta and N. Matsuzawa *et al.*, 2005. TDOA location system for IEEE 802.11b WLAN. Proceeding of the IEEE Wireless Communications and Networking Conference, Mar. 13-17, IEEE Xplore Press, USA., pp: 2338-2343. DOI: 10.1109/WCNC.2005.1424880