A Novel Indoor Positioning Method Based on Location Fingerprinting

Jiang Long Liu[1], Yi He Wan[2], Bao Gen Xu[2], Si Long Tang[2], Xue Ke Ding[2], Qun Wan[1]
[1]Department of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu, 611731, China
[2]Joint Lab of Array Signal Processing, TongFang Electronic Science and Technology Co. Ltd, Jiujiang, 332007, China
e-mail: byby-liu@163.com

***Abstract*—Indoor location based services (LBSs) is an emerging subject in recent years. To solve the problem of building the fingerprint database manually in the Location Fingerprinting (LF) methods, an analytical fingerprinting method is proposed in this paper. It can reduce the cost and has good properties on building database automatically and increasing the robustness against to environment. Moreover, the k nearest neighbor (k-NN) algorithm is used to search the closest fingerprint during the real-time localization phase. The simulation results are provided and the positioning performance is verified in comparison with that of classical methods. Finally, an experiment is carried out in an actual office room to verify the proposed method. Both the simulation and experiment results show that the new method is effective and can bring more accurate and robust localization substantially.**

I. Introduction

OVER the last several years, accurate, reliable and robust indoor locating protocols are increasingly required by more and more users and applications [1]. A real-time locating system gives the positions of mobile devices, and the position information can be used to many services such as navigation, tracking and monitoring [2]. As the positioning systems have been widely developed, real-time positioning and tracking features have been included in many mobile devices. For example, global positioning system (GPS) is the best known and the most widely used positioning system. However, there are shortcomings when GPS is used in the non-line-of-sight (NLOS) environments, such as built-up densely areas and indoor environments, due to the heavy influences of various noise sources, multipath effects and non-line-of-sight transmission [3].

IEEE 802.11Wireless Local Area Networks (WLANs) Wi-Fi has become widely installed. It is a good choice to use Wi-Fi signal strength to determine position as it is easy to be taken by a smartphone. Based on this idea, there are two kind of typical indoor positioning approaches, propagation-based approach and LF based approach. Propagation-based approaches [4-6], estimate the position by measuring the received signal strength (RSS) with path loss. However, propagation model cannot be established in common indoor environment because of the dynamic and unpredictable nature of radio channel, which is always the major challenge for accurate indoor positioning.

Instead of using the propagation model to describe the relationship between RSS and position, LF-based approaches are proposed to improve the accuracy of indoor positioning. The LF-based approaches, such as the [7-9] described, locate devices by comparing online RSS readings with offline observations. Various schemes about constructing fingerprint database were approached in recent years, but there are still some drawbacks. Firstly, the LF requires an initially survey with a very large training dataset. Secondly, the LF is very sensitive to signal fluctuation due to the changes of infrastructures and channel interference among access points (APs) leading to inaccurate positioning. In this paper, the analytical fingerprinting method is proposed to solve these problems and can have good properties on building database automatically and weakening dependency on the environment. Thus, the proposed method is a kind of improved LF-based method.

Considering the location estimation algorithm in LF-based approaches, one simple solution is the k-NN algorithm, which estimates the mobile user's location by computing the centroid of the k closest neighbors that have the smallest Euclidean distance to the online RSS reading [10]. Such a system is easy to implement but the estimation is not very accurate. Another solution to the LP approach is to solve the problem by a statistical method, in which the probability of each potential position is analyzed using the Bayesian theory and kernel functions [11], assuming that the RSS readings from different APs are independent at every time instant. However, an explicit formulation of RSS distribution is challenging and the independence may not hold in real environments. Meanwhile, these probabilistic based systems often have high computational complexity, which makes it difficult to run on mobile devices with limited power and small memory. Therefore, we adopt the k-NN algorithm to determine the mobile device's location.

Similar with the classical LP methods, the analytical fingerprinting method can be completed with two phases: offline phase and online phase [12]. The offline training phase is to build an analytical fingerprint database. In the online location determination phase, target's location is estimated by matching the real-time RSS information with the pre-measured database based on the Hanmming distance. As already mentioned, what is different from the classical method is that the analytical fingerprinting method can build a

978-1-4799-3051-7/13/$31.00 ©2013 IEEE

fingerprint database automatically. Besides, it enables the positioning system mitigate the measurement noises effectively.

The remainder of this paper is organized as follows. Section II explains the basic theories of new indoor positioning method. The complete mathematical presentation and the location process are given. In section III, the simulation model is built to examine the performance of the algorithm by MATLAB. What's more, an experiment is carried out in an actual office room to verify the proposed indoor positioning method. Section IV summarizes our work and presents the characteristics of the proposed indoor positioning method.

II. THEORY

As mentioned above, the analytical fingerprinting method is an improved LF method, so it is necessary to give a brief description for the LF method. The LF approach first requires the collection of data $\{(\mathbf{d}_i,\mathbf{C}_i), i=1,\ldots,M\}$, for M locations in an area, where $\mathbf{C}_i$ is the known location coordinate of the i-th measurement, and the vector $\mathbf{d}_i$ is the "fingerprint" of the location $\mathbf{C}_i$. When a new fingerprint $\mathbf{d}$ is derived from a transmitter at an unknown location $\mathbf{A}$, we can locate $\mathbf{A}$ by searching for the fingerprint $\mathbf{d}_i$ which is the closest to $\mathbf{d}$ and estimate the location with the corresponding $\mathbf{C}_i$. In the new proposed method, we will build a total different fingerprint database automatically, which is only made up with logic vectors.

A. The offline phase

The purpose of this phase is to generate a fingerprint database based on distances. The geographical coordinates of M reference points (RPs) are known. Assuming that there are N APs, and the representation of the overcomplete dictionary $\mathbf{D}$ is shown as below:

$$\mathbf{D}=[\mathbf{d}_1,\mathbf{d}_2\cdots\mathbf{d}_M]_{L\times M} \tag{1}$$

where the atom $\mathbf{d}_k(k=1,\cdots,M)$ is the fingerprint of every RP. $L=C_N^2$ is the length of every vector $\mathbf{d}_k$.

The fingerprint vector $\mathbf{d}_k$ of the k-th RP is calculated as follows. The Euclidean distances between the every RP and N APs can be calculated easily. The distances can be presented:

$$\mathbf{r}=\begin{pmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,N} \\ r_{2,1} & r_{2,2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ r_{M,1} & r_{M,2} & \cdots & r_{M,N} \end{pmatrix} \tag{2}$$

where $r_{k,i}$ is the distance from k-th $(k=1,\cdots,M)$ RP to i-th $(i=1,\cdots,N)$ AP. The rows of $\mathbf{r}$, $\mathbf{r}_k$ $(k=1,2,\cdots,M)$, represents the distance readings at each RP, which can be referred to as

$$\mathbf{r}_k=[r_{k,1},r_{k,2},\cdots,r_{k,N}], k=1,2,\cdots,M \tag{3}$$

The fingerprint of each RP is defined as

$$\mathbf{d}_k=[d_1,d_2,\cdots,d_L]^T,\quad L=C_N^2 \tag{4}$$

The method to generate the fingerprint vector $\mathbf{d}_k$ by using distance vector $\mathbf{r}_k$ is shown by the below equation,

$$\mathbf{d}_k(l)=\begin{cases} 1 & r_{ki}<r_{kj} \quad (i=1,\cdots,N, j=i+1,\cdots,N) \\ & k=1,\cdots,M \quad l=(i-1)N-\dfrac{i(i+1)}{2}+j \\ 0 & r_{ki}\ge r_{kj} \quad (i=1,\cdots,N, j=i+1,\cdots,N) \end{cases} \tag{5}$$

Thus, the fingerprint vector $\mathbf{d}_k$ made up of binary number is called analytical fingerprint in this paper.

What need to be noticed is that many different locations have a same fingerprint, so these RPs need to be gathered and form a fuzzy region. The median point of the fuzzy region need to be calculated and is used to represent the all RPs inside the fuzzy region.

Suppose there are P kinds of different atoms in $\mathbf{D}=[\mathbf{d}_1,\mathbf{d}_2\cdots\mathbf{d}_M]_{L\times M}$, the variable q represents the total number of RPs in every fuzzy region. Then the median point can be calculated using the following equation,

$$\begin{cases} x_j=\dfrac{1}{q}\sum_{i=1}^{q}x_i \\ y_j=\dfrac{1}{q}\sum_{i=1}^{q}y_i \end{cases} \quad (j=1,2,\cdots P) \tag{6}$$

So far, there are P kinds of separated atoms and P geographical coordinates, which are the final fingerprint database and can be represented as $\{(\mathbf{d}_i,\mathbf{C}_i), i=1,\ldots,P\}$ and $\mathbf{C}_i=(x_i,y_i), i=1,2,\cdots P$.

B. The online phase

Assuming that there is a mobile device receives signals from all APs. By measuring signal strength, a series of RSSs can be obtained and represented as $[rss_1,rss_2,\cdots rss_N]$. RSS sort sequence can be generated by pairwise comparisons of RSSs.

$$RSS_l=\begin{cases} 1 & rssi_i>rssi_j \quad (i=1,\ldots,N, j=i+1,\ldots,N) \\ & l=(i-1)N-\dfrac{i(i+1)}{2}+j \\ 0 & rssi_i\le rssi_j \quad (i=1,\ldots,N, j=i+1,\ldots,N) \end{cases} \tag{7}$$

According to (7), the RSS binary vector can be represented a $\mathbf{RSS} = [RSS_1, RSS_2, RSS_3, ..., RSS_L]^T, L = C_N^2$ .In accordance with the relationship between RSS and distance, we know that $\mathbf{RSS}$ and one of $\mathbf{D}$ are matching only when the target is in this fuzzy region.

We use the k-NN algorithm based on the Hamming distance to search the k closest fingerprints and determine its location. The Hamming distance is the total number of the corresponding different symbols between two different vectors [13]. The Hamming distances between the $\mathbf{RSS}$ with all the analytical fingerprints can be calculated by,

$$\mathbf{hm}(j) = \left\| \mathbf{RSS} \oplus \mathbf{d}_j \right\| \qquad j = 1,2,3,\cdots P \tag{8}$$

where $\oplus$ denotes XOR operator. The smaller $\mathbf{hm}(k)$ is, the more similar $\mathbf{RSS}$ and $\mathbf{d}_j$ are. In this way, we choose the k most similar fingerprints and calculate the median point of these fuzzy regions, in the same way as equation (7). Then, the median point is selected as the estimated point. So far, positioning process is accomplished.

III. Simulation and Experiment Results

A. Before Simulation results

The simulation model to examine the performance of the algorithm can be made in this section using MATLAB. Discussion is started from the ideal situation which is the simplest. At first, an off-line fingerprint database is built according to the locations of APs. Five APs are fixed in an area of $4m \times 4m$ where 6561 mesh nodes are RPs. Then, the distribution of fuzzy regions and their median points can be established easily, as shown in Fig.1.

Based on the model, the positioning program is run for 6561 times. The error distance between position estimated coordinate and the original coordinate is calculated and collected under same signal-to-noise ratio (SNR) each time. Positioning accuracy of the algorithm can be obtained by processing the simulation results with statistical perspective [12].

In order to verify the performance of the new method, the analytical fingerprinting algorithm is compared with other methods. The same data is processed by the Maximum Likelihood (ML) multilateration method, which obtains the Minimum Mean Square Estimate (MMSE) of a node's position from a set of noisy distance measurements [14-16].

According to Fig.2, The new method performs better and achieves an accuracy of 1m with 47% probability, while that of ML multilateration method is just up to 27.5% when SNR is 5dB. In addition, the average error distance of the new positioning algorithm is 0.2486m in noiseless situation.

B. Experiment results

Beside the simulation results above, an experiment in an actual indoor environment is carried out to give a further verification to the proposed method. Our experiment is

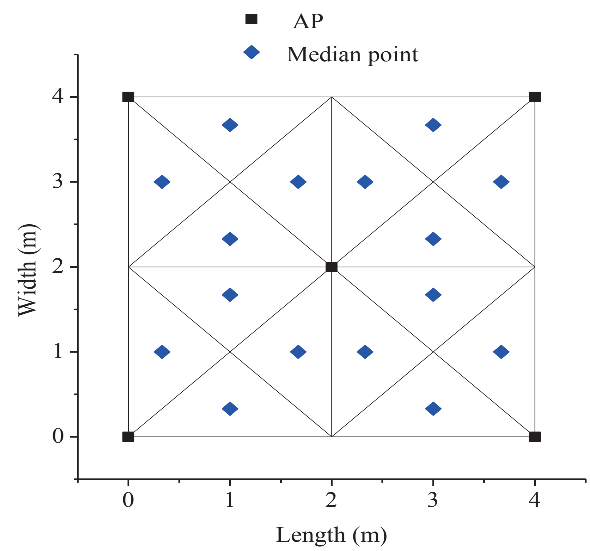

Fig. 1. The simulation setups and distribution of fuzzy regions.

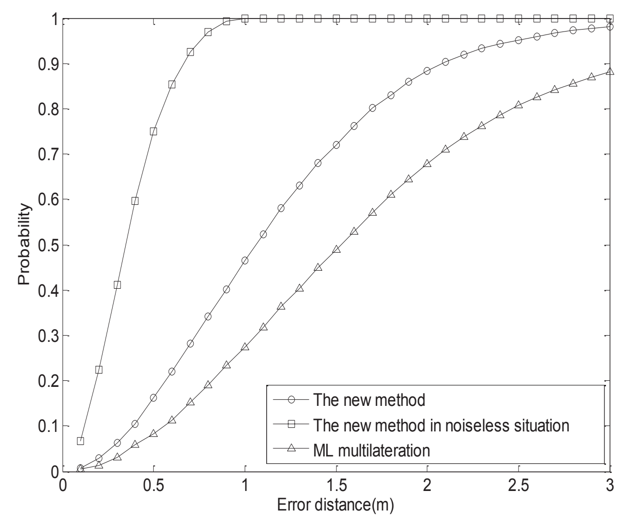

Fig. 2. The comparative performance of the new method and classical method when SNR=5dB and the new method in ideal condition.

implemented in the office area of an electronic company which has a size of $41.35m \times 20.45m$. The top view of the area is shown in Fig. 3. In this indoor area, we deployed 5 APs of Wi-Fi (the green dots in Fig. 3) and a mobile device. 26 sample points are also fixed whose locations are also shown in Fig. 3.

According to the locations of APs, the off-line database can be calculated by a computer in advance. The fuzzy regions distribution of this database, as show in Fig. 4, is not uniform design and regular. In our experiment, this indoor environment is divided into 872×409 pixels which each one pixel represents an area of $0.05m \times 0.05m$. By the real-time positioning at sampling points, the error distance is described by an arrow from the actual location to the estimated location as shown in Fig.5. According to the scale of map, the average error distance is within 2m and even error distances of some locations can be less than 0.17m.

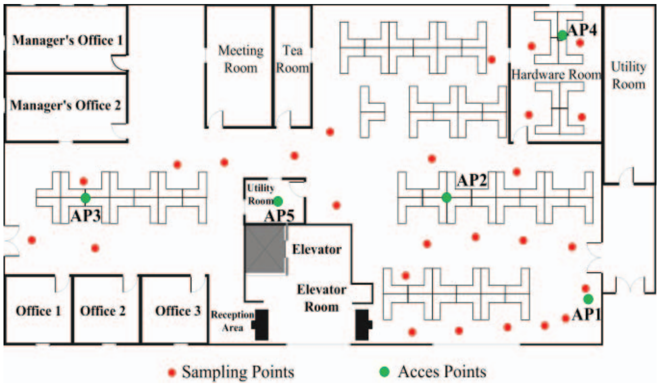

Fig. 3. The top view of the indoor experiment environment.

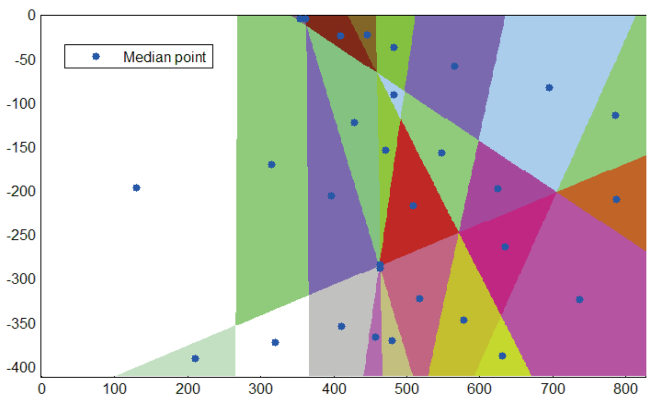

Fig. 4. The corresponding fuzzy regions distribution of the indoor experiment environment.

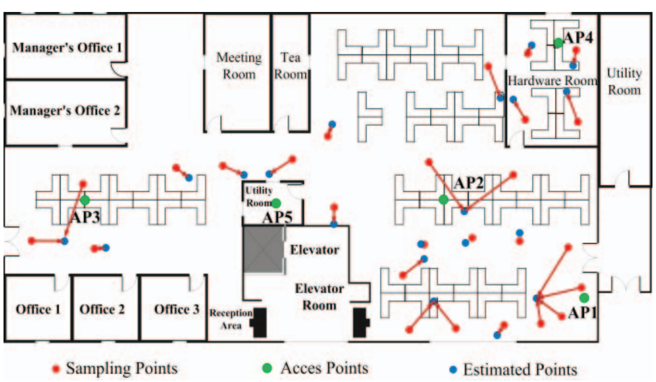

Fig. 5. The positioning performance at sampling points. The arrow denotes the error distance from the actual location (sampling point) to the estimated location.

IV. Conclusion

A novel indoor positioning method based on analytical fingerprinting is proposed in this paper. Simulation and experiments are carried out to verify the advantages of this proposed analytical fingerprinting based method. Firstly, analytical fingerprinting database can be built up automatically and the cost can be decreased. Secondly, analytical fingerprinting can be used to achieve higher probability of identification and more accurate position. In addition, not only the simulation results but also a series of measurement data verified that it enables the positioning system mitigate the measurement noises effectively.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under grant 61172140.

References

[1] Vossiek M, Wiebking L, Gulden P, Wieghardt J, Hoffmann C, Heide P , "Wireless local positioning," *Microwave Magazine, IEEE*, vol.4, no.4, pp. 77- 86, Dec. 2003.

[2] Flora C, Ficco M, Russo S, Vecchio V , "Indoor and outdoor location based services for portable wireless devices," *25th IEEE International Conference on Distributed Computing Systems Workshops*, pp. 244-250, 6-10 June 2005.

[3] B. Hofmann, H. Wellinhof and H. Lichtenegger, "GPS: Theory and Practice", Springer-Verlag, Vienna, 1997.

[4] R. Jan and Y. Lee, "An indoor geolocation system for wireless LANs," *Int. Conference, Parallel Processing Workshops*, pp.29-34,2003.

[5] P. Prasithsangaree, P. Krishnamurthy, and P. Chrysanthis, "On indoor position location with wireless LANs," *The 13th IEEE International Symposium, Personal, Indoor and Mobile Radio Communications*.,vol. 2, 2002.

[6] J. Kwon, B. Dundar, and P. Varaiya, "Hybrid algorithm for indoor positioning using wireless LAN," *IEEE 60th Vehicular Technology Conference*, VTC, vol. 7, 2004.

[7] B. Li, Y. Wang, H. Lee, A. Dempster, and C. Rizos, "Method for yielding a database of location fingerprints in WLAN," *Communications, IEE Proceedings*, vol. 152, no. 5, pp. 580- 586, 2005.

[8] N. Swangmuang and P. Krishnamurthy, "Location Fingerprint Analyses Toward Efficient Indoor Positioning," *Sixth Annual IEEE International Conference on Pervasive Computing and Communications*, pp. 101- 109, 2008.

[9] S. Fang, T. Lin, and P. Lin, "Location Fingerprinting In A Decorrelated Space," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 20, no. 5, pp. 685- 691, 2008.

[10] J. Ma, X. Li, X. Tao, and J. Lu, "Cluster Filtered KNN: A WLAN Based Indoor Positioning Scheme," *Proc. Int' l Symp. World of Wireless, Mobile and Multimedia Networks*, pp. 1-8, June 2008.

[11] R. Singh, L. Macchi, C. Regazzoni, and K. Plataniotis, "A Statistical Modelling Based Location Determination Method Using Fusion in WLAN," *Proc. Int' l Workshop Wireless Ad-Hoc Networks*, 2005.

[12] Yanying Gu, Lo A, Niemegeers I, "A survey of indoor positioning systems for wireless personal networks," *Communications Surveys & Tutorials, IEEE* , vol.11, no.1, pp.13-32, First Quarter 2009.

[13] Steane, Andrew M. "Error correcting codes in quantum theory." *Physical Review Letters*, vol.77, no.5, pp.793, 1996.

[14] Andreas Savvides, Chih-Chieh Han, Mani B. Strivastava, "Dynamic Fine-Grained Localization in Ad-Hoc Networks of Sensors," *Seventh Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 2001)*, Rome, Italy, July 16-21, 2001.

[15] Yoshikazu Ohta, Masashi Sugano, Masayuki Murata, "Autonomous Localization Method in Wireless Sensor Networks," *Pervasive Computing and Communications Workshops, 2005, Third IEEE International Conference*, pp.379-378, 8-12 March 2005.

[16] Cheung K.W, So H.C, Ma W.-K Chan Y.T , "Least squares algorithms for time-of-arrival-based mobile location," *IEEE Transactions on Signal Processing*, vol.52, no.4, pp. 1121-1130, April 2004.