Crowdsensing-based Organic Fingerprint for Wi-Fi Localization

Wenzheng Gao, Ling Pei*, Changqing Xu,Peilin Liu
Shanghai Key Laboratory of Navigation and Location-based Services
School of Electronic Information and Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China
*Corresponding author: ling.pei@sjtu.edu.cn

***Abstract*—In a Wi-Fi location system, the accurate fingerprint database is the key to ensure accurate positioning. Several methods have been proposed for database generating and database training, but most of them have been trying to get a static fingerprint database which is only available in stable signal environment. In this paper, we proposed a crowdsensing-based fingerprint updating algorithm which makes the fingerprint organic and guarantee the positioning accuracy both in stable and unstable Wi-Fi signal environments. To get this, we first analyze the Wi-Fi signal environments, the instability of Wi-Fi signal environments will deteriorate the positioning performance of crowdsensing-based Wi-Fi location system if the fingerprint database is fixed. To do fingerprint updating, we first proposed a fingerprint updating algorithm based on the Euclidean distance between distinct fingerprints, which effectively reduce the importing error in crowdsensing-based database generating and signal sampling. The algorithm works well in stable and short-term Wi-Fi environments, but it can't solve the problem when tremendous changes occur in Wi-Fi environments. In order to solve the problem, we proposed an advanced algorithm based on the first one: Fingerprint updating based on Wi-Fi fingerprints' reliability model. In this algorithm, we assign every fingerprint a new property, reliability, which decays with time. By utilizing the reliability of the fingerprint, the location system can automatically detect the change of Wi-Fi environments and remove the outdated fingerprints. With this updating algorithm, we can get an organic fingerprint database.**

Keywords—Wi-Fi Localization;fingerprint database;fingerprint updating;reliability model

I. INTRODUCTION

Indoor navigation has been widely concerned since the last decade. Compared with outdoor navigation, indoor navigation faces a series of challenges because of the much more complex environment, where the Global Navigation Satellite Systems (GNSS) do not work. With the rapid development of Wireless Fidelity (Wi-Fi), the penetration of Wireless Local Area Networks (WLAN) basic installation has reached an unprecedented high level, and fingerprint-based Wi-Fi localization has been well accepted to solve the indoor localization in recent years. The popularity of smartphones which can easily get the Wi-Fi signal strength makes it possible to satisfy most of our needs of indoor localization [1]-[5].

Two techniques have been investigated using Wi-Fi signals for indoor positioning, both are based on the measurement of signal strengths of the device to be located. The first, trilateration, which attempts to convert the measurements into distances between the device and the Wi-Fi access point (AP). The second technique, fingerprint-based positioning. It first requires the construction of the fingerprint database, which record the signal strengths from different APs at different reference points in the desired coverage area. The location of the device is then obtained by measuring the signal strengths at its location, and comparing it with different reference fingerprints in the database [12].

Owing to the complexity of indoor environments and the underlying nature of Wi-Fi signals, it is difficult to convert signal strength measurements into accurate distances in trilateration positioning. So the trilateration is hardly used commercially. The main disadvantage of fingerprint-based positioning is the labor cost of generating the database before real-time positioning. Conventianal fingerprint-based positioning system requires an extensive site survey to guarantee the accuracy of the location of the reference points, so the workload of generating a fingerprint database is heavy [10]-[12]. However, the advantage of accuracy makes it a better choice in most cases. Recent research of fingerprint-based positioning has been trying to utilize inertial sensors or magnetic sensors as assistances for getting the accurate location of reference points, which can reduce the workload and generate the fingerprint database under the condition of semi-supervised or unsupervised [6]-[9]. Another important cost is the one that associated with maintaining the fingerprint database over time. As the signal environment of Wi-Fi is unstable, both the aging and the replacement of APs will change the correlations between locations and Wi-Fi signal strengths. Therefore, for accurate positioning results, Wi-Fi location requires the support of an organic fingerprint database which can be updated by using the crowdsensing data.

Most research in Wi-Fi positioning is currently focused on real-time positioning algorithm or fingerprint generating to improve the positioning accuracy, efficient fingerprint updating algorithms are infrequent in Wi-Fi location systems [12]. Such a fingerprint updating algorithm is presented in this paper. It relies on the crowdsensing data contributed by all the users in the coverage area, which can be utilized to generate the initial raw fingerprint database and update the database over time. The paper is organized as follows. In Section II, we present the sources, the structure and the application method

978-1-5090-2879-5/16/$31.00©2016 IEEE

of fingerprint database. We propose our updating algorithms in Section III. The results of our experiments are presented in Section IV, followed by the conclusions.

II. DATABASE OF FINGERPRINT-BASED POSITIONING

In this section, we will show the structure and properties of fingerprint database in the Wi-Fi location system. The sampling, generation of the database and how the fingerprint database works in real-time positioning phase will be presented in order. At last, we will analyze the necessity of database maintaining and updating.

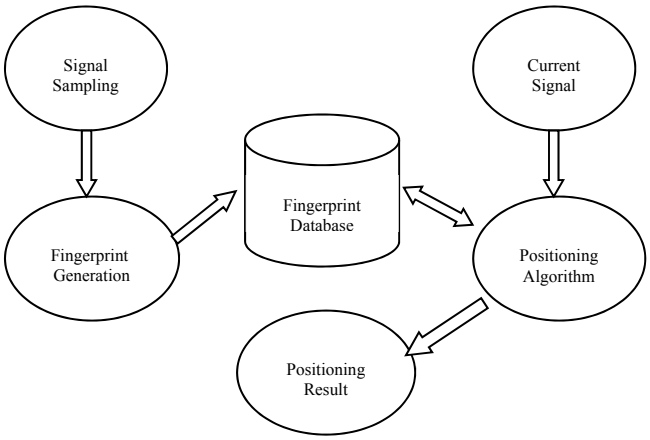


Fig. 1. Flow diagram of fingerprint-based location system

Figure 1 shows the flow of fingerprint-based location system. We first lay out the mathematical framework for the approach, then give details about each component of the system.

In conventional signal sampling method, we need a skilled surveyor to stand right at the reference points and sample the Wi-Fi signals in stillness for one minute or more. The strict constraints increase the workload and make the database difficult to be rebuilt. In our method, Pedestrian dead reckoning (PDR) is utilized as an assistance for getting the location of the reference points. PDR can detect the signals of inertial sensors and calculate the location of the device in real time. With the PDR running, the device sample the Wi-Fi signals synchronously. So the sampling result can be formulated as:

$$fingerprint_n = \left\{ \begin{matrix} (x_n, y_n, z_n) \\ (MAC_1, RSS_{n1}), (MAC_2, RSS_{n2}) \dots (MAC_m, RSS_{nm}) \end{matrix} \right\} \quad (1)$$

The inertial sensors PDR required and Wi-Fi sensors are all possessed by smartphones, so all the users with smartphones can contribute to signal sampling under the condition of unsupervised. The rich sources of crowdsensing data make the fingerprint updating available.

After signal sampling, we get series of data combined with location and Wi-Fi signal strengths. However, the data can't be utilized in database generation directly, because there are too many locations in the data. If we use all the locations as reference points just like in conventional methods, the database size will steadily increase as the sampling data are processed, which may reduce the efficiency of fingerprint database retrieval and affect the real-time capability of the location system.

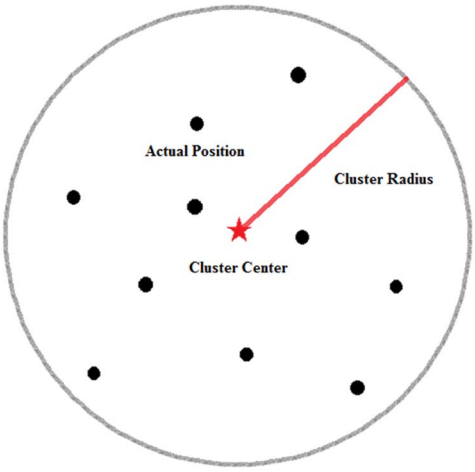


Fig. 2. Schematic of clustering

In our method, we choose some locations in the coverage area as reference points, and do clustering to the sampling data. As Figure 2 shows, the sampling data with similar location are clustered to the nearest reference point.

When the fingerprint data of one reference point is enough, we calculate the distribution probabilities of the Received Signal Strength (RSS) values from -30dBm to -90dBm from each AP. The RSS value is considered as Gaussian distribution from one AP at each reference point [10]. In (2), x is the RSS value, μ is the mean value of Gaussian distribution, σ is the standard deviation of Gaussian distribution.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (2)$$

		-90dBm	-89dBm	-88dBm …
Reference Point1	AP_1_MAC	Prob11	Prob21	Prob31 …
	AP_2_MAC	Prob12	Prob22	Prob32 …
	AP_3_MAC …	Prob13 …	Prob23 …	Prob33 … …
Reference Point2	AP_1_MAC	Prob11	Prob21	Prob31 …
	AP_2_MAC	Prob12	Prob22	Prob32 …
	AP_3_MAC …	Prob13 …	Prob23 …	Prob33 … …

Fig. 3. Format of fingerprint database

The sampling Wi-Fi signal strengths are continuous values, we discretize the values for ease of recording. The

probability of the RSS value from -30dBm to -90dBm can be considered as the integral of the probability distribution function from -0.5 to +0.5.

$$p_{\mathrm{rss}} = \int_{\mathrm{rss}-0.5}^{\mathrm{rss}+0.5} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-u)^2}{2\sigma^2}\right) dx \quad (3)$$

Thus, the fingerprint database in the format of Figure 3 is generated.

In the real-time positioning phase, we use Maximum Likelihood Estimate (MLE) to get users' location. When the user is in the coverage area, we will receive a RSS vector from his smartphone, $S = \{RSS_1, RSS_2, \ldots RSS_n\}$ which is sampled at the user's location. This vector record the Wi-Fi signal strengths from all the APs and will be used in fingerprint matching.

The probability of an RSS value from the i-th AP at a given reference point stored in the database is $p(RSS_i|RP_n)$. Since the probability distribution of RSS value from different AP is independent, the probability of S in the reference point can be calculated by the accumulated probability:

$$p(S|RP_n) = \prod_{i=1}^{N} p(RSS_i|RP_n) \quad (4)$$

According to the Bayesian formula, the probability that the user is right at the i-th reference point when the received vector is S can be considered as:

$$p(RP_i|S) = \frac{p(S|RP_i)}{\sum_{j=1}^{M} p(S|RP_j)} \quad (5)$$

According to MLE, after picking out K reference points with the highest probabilities, the user's location is estimated as the weighted sum of the locations of the reference points:

$$L = \sum_{i=1}^{K} p(RP_i|S)\, RP_i \quad (6)$$

As is shown above, the real-time positioning in fingerprint-based Wi-Fi location system is a kind of fingerprint matching process. MLE will match the user's location to the reference points which are similar to the user's location in Wi-Fi signal strengths. So the accuracy of the fingerprint database directly affects the positioning performance of Wi-Fi location system.

Figure 4 shows the Wi-Fi signal strengths' variation at the same location. We can see that the Wi-Fi signal strength is unstable. Even though the APs and smartphones are all normally working, the sheltering of pedestrians or the daily change of air humidity will cause the temporary small-scale change of Wi-Fi signal strengths. The positioning performance will be affected accordingly. However, these small-scale changes are reversible, in other words, the location system will be back to normal after a while.

Except for the temporary changes, there are also many permanent changes in the indoor environments. The move of the furniture may increase or decrease the sheltering between the APs and the users' smartphones; the aging of the wireless routers (AP) will cause the attenuation of the received strengths; the shutting off of the wireless routers will set all the RSS from the AP to -120dbm.

Figure 5 shows the positioning performance of a location system without database maintaining for one year. Most of the positioning results are far away from the ground truth. The location system hardly gives the right positioning result unless the database is rebuilt. As is mentioned above, the reconstruction of fingerprint database requires huge workload and keep positioning unavailable for quite a long time. So the organic fingerprint database is significantly necessary.

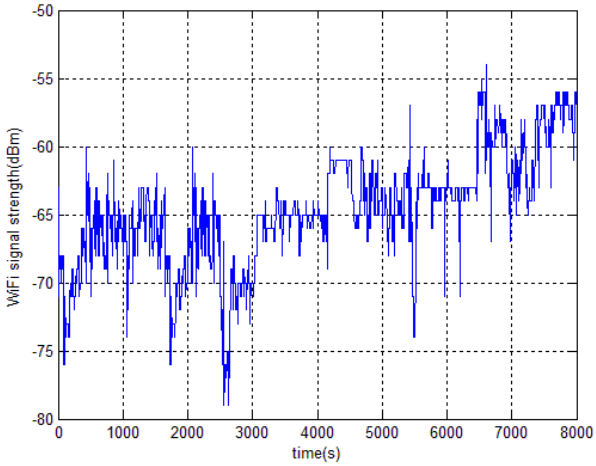


Fig. 4. Wi-Fi signal strength over time

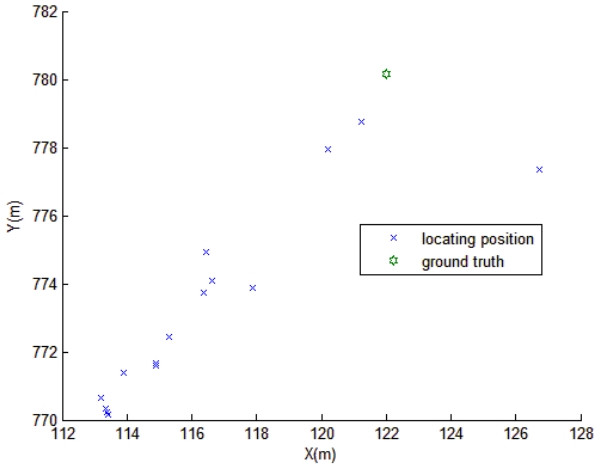


Fig. 5. The positioning performance of location system with out-of-date database

III. FINGERPRINT DATABASE UPDATING

In Section II, a fingerprint-based Wi-Fi location system was presented, the maintaining and updating of the fingerprint database was proved necessary in the location system. In this Section, our fingerprint updating method will be presented, which can detect the change of the Wi-Fi environment and update the fingerprint automatically.

A. Fingerprint updating based on the similarity between fingerprints

As is mentioned above, the positioning result of the location system is calculated with MLE:

$$L = \sum_{i=1}^{N} p_i \times RP_i \quad (7)$$

In (7), p_i is the likelihood probability that the user is at the

i-th reference point. However, some positioning errors may happen as is shown in Figure 6. In fingerprint matching, we found several matched fingerprints from the database because they are similar with the signal sampled by the smartphone. The matched fingerprints were clustered to the nearest reference points in database generation, but the fingerprints' real locations may be far from the reference points due to the accumulative error of PDR or the inappropriate choice of the clustering radius.

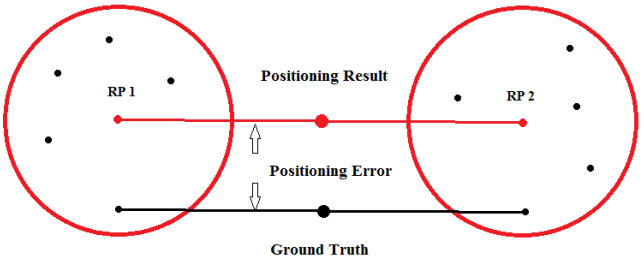


Fig. 6. Positioning error in the location system

The positioning error in Figure 6 will never happen in conventional location system with supervised database generation, because all the fingerprints in the database are sampled right at the reference points. In Wi-Fi location system, a fingerprint can be shown as (8), the sampled signal vector in real-time positioning phase can also be seen as a fingerprint. The similarity can be measured as (9), which means two fingerprints are similar when the signal Euclidean distance between them is small enough.

$$S_n = (RSS_{n1}, RSS_{n2}, \ldots RSS_{nm}) \tag{8}$$

$$\mathrm{Sim}_{ij} \propto \frac{1}{|S_i - S_j|} \tag{9}$$

According to the propagation model of Wi-Fi signal in (10) and (11), when the gains of the transmitting end and receiving end are fixed, the RSS is correlated with the transmitting signal strength and the distance between the AP and the user. In most Wi-Fi environments, the transmitting signal strengths of the APs are stable. So the RSS is only correlated with the distance between the AP and the user. As is shown in Figure 7, the user's location can be calculated with 3 APs at least. So the similarity between fingerprints should be equivalent to the similarity between locations. However, utilization of PDR and fingerprints clustering may disturb the equivalence relation between fingerprints' similarity and locations' similarity.

$$P_r = P_t \times G_t \times G_r \times \left(\frac{\lambda}{4\pi d}\right)^2 \tag{10}$$

$$RSS_i(dbm) = 30 + 10\log_{10} P_r(W) \tag{11}$$

In our method, we use the updating based on fingerprints' similarity to process the fingerprints in the original fingerprint database and the ones in the updating data. The updating consists of two steps: Acquiring the standard fingerprints and High-pass filtering based on fingerprints' similarity.

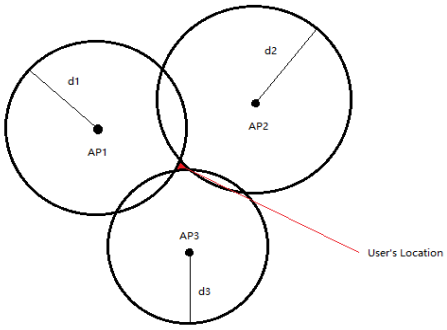


Fig. 7. Schematic of user's location confirmed by 3 APs

Acquiring the standard fingerprints: In order to remove the fingerprints which are far from the corresponding reference points in location, we must find the distinctions of these fingerprints. The RSS value from each AP is considered as a Gaussian distribution at the reference point. When the user is away from the reference point, the RSS value from the APs will differ from the one sampled right at the reference point, which increase the Euclidean distance between the two fingerprints. Therefore, we can conclude that the fingerprints far from the reference points in location are far from them in fingerprints' Euclidean distance.

However, due to the accumulated error, PDR can't guarantee the locations' accuracy in signal sampling phase. It's difficult to find the fingerprints sampled right at the reference points with PDR's results. We will use the statistical properties of the fingerprints in the database and updating data to find the standard fingerprint of each reference point. In database generation phase, locations which are close to the reference points are clustered. The locations are considered Gaussian distributed to the reference point. That means the average value of the locations will be the reference point with sufficient samples. With the equivalence relation between fingerprints' similarity and locations' similarity we have proved above, the average value of the fingerprints can be considered the fingerprint of the reference point.

When the fingerprints in the database are $S_m = (RSS_{m1}, RSS_{m2}, \ldots, RSS_{mN})$, and the fingerprints in the updating data are $S_k = (RSS_{k1}, RSS_{k2}, \ldots, RSS_{kN})$, the standard fingerprint is:

$$S_s = \frac{\sum_{i=1}^{M} S_m + \sum_{i=1}^{K} S_k}{M+K} \tag{12}$$

High-pass filtering based on fingerprints' similarity: After acquiring the standard fingerprint of each reference point, the fingerprints' similarity is given by (13). For each reference point, we can sort the fingerprints in the original

database and updating data by their similarity to the standard fingerprint. M fingerprints with the highest similarity will be saved as the new fingerprint database of the reference point. The fingerprints with lower similarity will be filtered out. According to the relationship between fingerprints' similarity and locations' similarity we have mentioned above, the fingerprints which are far from the standard fingerprints in Euclidean distance are far from the cluster center, so they should be eliminated from the database. The effect of high-pass filtering is shown in Figure 8.

$$Sim_i = \alpha \times (|S_i - S_s|)^{-1} \quad (13)$$

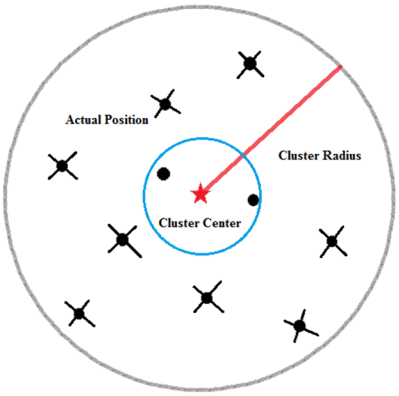


Fig. 8. Schematic of the effect of high-pass filtering

The flow diagram of our updating process is shown in Figure 9. Figure 10 shows the practical positioning performance of this updating method. In Figure 10 (a), the positioning results locate dispersedly around the ground truth, that's because there are some bad fingerprints with location errors, when the signal sampled is matched to a bad fingerprint, the positioning error occurs. In Figure 10 (b), the positioning results obviously converge, but more updating data are required to amend the database. In Figure 10 (c), most of the positioning results locate near the ground truth.

However, the updating method may break down in such a case: When the amount of updating data is smaller than the database, the change of the Wi-Fi environment may be ignored. The acquiring of the standard fingerprint is the core of this updating method. In (12), we can know that the standard fingerprint represents the Wi-Fi signal characteristic of the reference point only if the fingerprints in the updating data are more than the ones in the database. However, in unsupervised signal sampling, the condition can't be guaranteed. So when the Wi-Fi environment changes, the standard fingerprint may be wrong owing to the insufficiency of updating data. Figure 11 shows the positioning performance in this case. More position errors occur without the right database updating.

The insufficiency of updating data is not a coincidence in unsupervised signal sampling. As a matter of fact, it often happens in the Wi-Fi location system. Therefore, a more robust updating method is required.

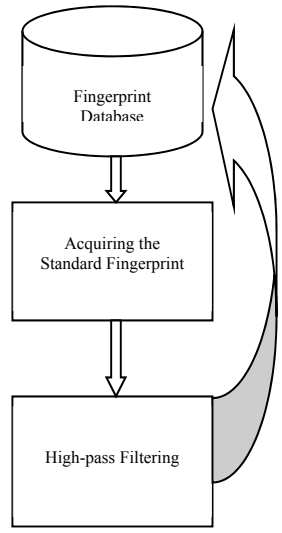
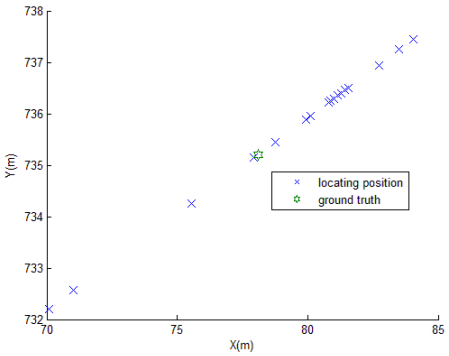
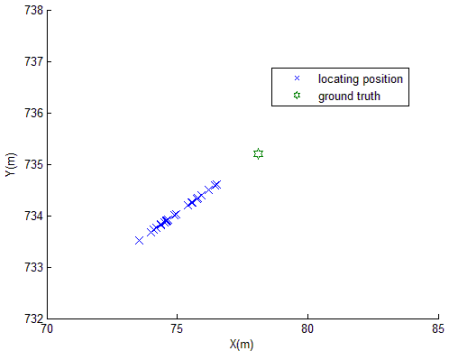


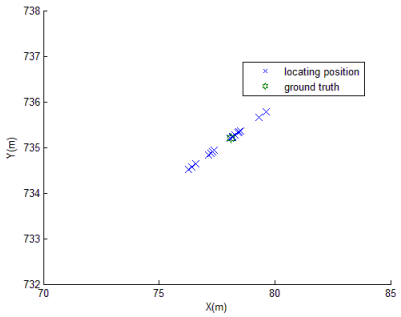
Fig. 9. Flow diagram of updating based on the similarity between fingerprints



(a)



(b)



(c)

Fig. 10. Positioning performance of updating based on the similarity between fingerprints (a) Initial database (b) Database of the first updating (c) Database of the second updating

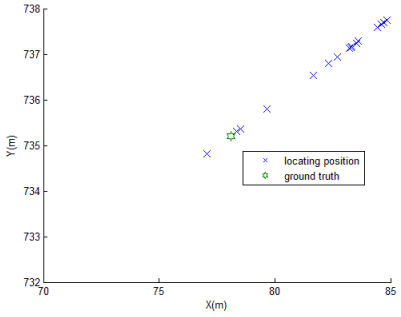


Fig. 11. Positioning performance of updating in unstable Wi-Fi environment

B. Fingerprint updating based on Wi-Fi fingerprints' reliability model

In order to make the fingerprint updating more robust in unstable Wi-Fi environments, we propose a new updating algorithm based on the updating method above: Fingerprint updating based on Wi-Fi fingerprints' reliability model.

The complexity of indoor environment makes the Wi-Fi signal strength difficult to be confirmed by the propagation model. There are many factors which may disturb Wi-Fi signals' stability, such as the sheltering of furniture, aging of the APs and the replacement of APs. All these factors will affect the Wi-Fi signal strengths complexly and make it impossible to measure the changes of the Wi-Fi environment by an accurate mathematical model. Therefore, we introduce a new property of the fingerprint: reliability. In our algorithm, a fingerprint's reliability is only correlated with the time when the fingerprint is sampled. The principle of the reliability model is: The newly sampled fingerprint is always more reliable than the old fingerprint in representing the signal property of the present.

As is shown in (12), when the standard fingerprint is calculated, the fingerprints in the updating data own the same weights as the ones in the original database. When the amount of the fingerprints in the updating data is much bigger than the one in the database, the updating data dominate the standard fingerprint by the strength in numbers. The changes of the Wi-Fi environment will be detected because they are recorded in the updating data. On the contrary, when the amount of the fingerprints in the updating data is smaller than the one in the database, the newly sampled fingerprints may not be elected to the updated database because the changes of Wi-Fi environment make them different from the fingerprints in the original database. So the changes are ignored.

We will give the fingerprints different weights according to the reliability model. The number of the APs is N, the fingerprints in the original database and the updating data can be represented as $S = (RSS_1, RSS_2, \dots, RSS_N)$. We set the number of the fingerprints in the database as M, and the number of the fingerprints in the updating data is K which is not known previously. The fingerprints in the updating data have the weights which are T times more than the ones in the original database. The weighted average value of all the fingerprints is:

$$S_w = \frac{T\times\sum_{i=1}^{K} S_k + \sum_{i=1}^{M} S_m}{T\times K + M} \tag{14}$$

For example, if the first AP breaks down, the Wi-Fi signal strengths from this AP will be added by an offset Δ. In (14), the RSS value from other APs is not affected. The weighted average value of the signal strengths from the first AP is:

$$\overline{RSS_1'} = \frac{TK(\overline{RSS_1}+\Delta)+M\overline{RSS_1}}{TK+M} = \overline{RSS_1} + \frac{TK\Delta}{TK+M} \tag{15}$$

The Euclidean distance between the fingerprint and S_w is:

$$\text{Edistance} = \sqrt{\left(RSS_1 - \overline{RSS_1'}\right)^2 + \left(RSS_2 - \overline{RSS_2'}\right)^2 + \cdots + \left(RSS_N - \overline{RSS_N'}\right)^2} \tag{16}$$

Because the latter N-1 APs aren't affected, their corresponding variables in (16) can be omitted, the most noticeable difference between the updating data and the original data is in $\left|RSS_1 - \overline{RSS_1'}\right|$. For the fingerprints in the original database, $\left(RSS_1 - \overline{RSS_1'}\right)$ is considered as a Gaussian distribution with a mean of $-\frac{TK\Delta}{TK+M}$ and a variance of σ when the Wi-Fi signal variance is σ. For the fingerprints in the updating data, $\left(RSS_1 - \overline{RSS_1'}\right)$ is $\Delta - \frac{TK\Delta}{TK+M}$. If we want the fingerprints in the updating data to be elected as the fingerprints in the new database, the Euclidean distance of them should be smaller than at least K/M of the fingerprints in the original database when K is smaller than M. The restricted condition of T can be given:

$$T \geq \frac{M\left(1-\frac{\phi^{-1}(1-\frac{K}{2M})\times\sigma}{\Delta}\right)}{K\left(1+\frac{\phi^{-1}(1-\frac{K}{2M})\times\sigma}{\Delta}\right)} \tag{17}$$

In (17), $\phi^{-1}(x)$ is the inverse function of Normal Distribution Function. In the reliability model, the newly sampled fingerprint owns a bigger weight than the older one. The attenuation characteristic can be presented by the Exponential Function $f(x) = e^{-\alpha x}$. When the updating interval is set to Δt, the parameter α is:

$$\alpha = \frac{\ln T}{\Delta t} \tag{18}$$

$$R(t) = \begin{cases} e^{-\frac{lnT}{\Delta t}(t_0 - t)} & t_0 - t < t_m \\ 0 & t_0 - t > t_m \end{cases} \quad (19)$$

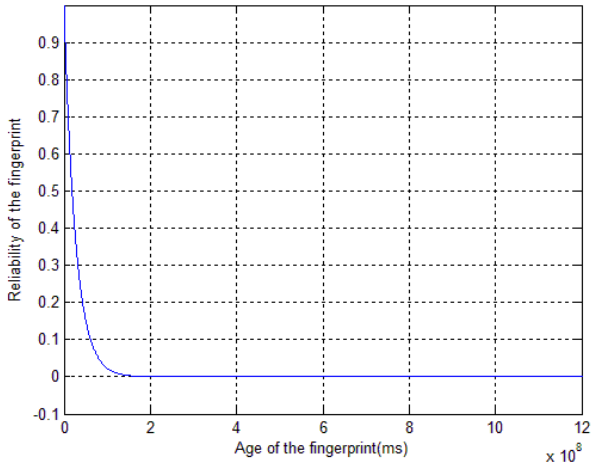


Fig. 12. Reliability model of fingerprints

In (19), the reliability model is shown, t_0 is the current time, t is the time when the fingerprint is sampled, t_m is the threshold of outdated fingerprints. When a fingerprint is judged outdated, we will set its reliability to 0 to reduce the computational complexity of fingerprint updating.

The updating based on Wi-Fi fingerprints' reliability model consists of 3 steps: Acquiring the standard fingerprints, High-pass filtering based on fingerprints' similarity and Dynamic setting of AP.

Acquiring the standard fingerprints: The standard fingerprint of the reference point is the weighted average value of the fingerprints in the updating data and the database.

$$S_s = \frac{\sum_{i=1}^{K} R(t_i) \times S_i + \sum_{i=1}^{M} R(t_i) \times S_i}{\sum_{i=1}^{K} R(t_i) + \sum_{i=1}^{M} R(t_i)} \quad (20)$$

According to the reliability model above, when K is much bigger than M, the standard fingerprint is more similar to the fingerprints in the updating data, which can help the database to detect the change of the Wi-Fi environment. When K is smaller than M, although the fingerprints in the original database have the strength in numbers, the newly sampled fingerprints are assigned bigger weights, which will make the standard fingerprint more similar to the state of the present.

High-pass filtering based on fingerprints' similarity: The high-pass filtering in this method is same as the one above. However, owing to the reliability model, some "abnormal" fingerprints will be elected as the fingerprints in the new database. Some of them are caused by the changes of the Wi-Fi environment. Others are caused by the accumulated errors of PDR. The former fingerprints will be repeated to confirm the changes, the latter fingerprints will not be repeated frequently so they will be filtered out later.

Dynamic setting of AP: In the conventional Wi-Fi location system, the choices of APs are fixed which may reduce the robustness of the system. As the replacement of APs happens over time, there are fewer and fewer valid APs in the fixed AP list and the system is running a higher and higher risk of breaking down.

Such a dynamic setting of AP is proposed: Initially, we will select 3M alternative APs which appear most frequently in the crowdsensing data. Then we choose the first M APs from the alternative list as the APs for generating the initial database. The alternative list may change after analyzing the crowdsensing data every time before the updating. After the updating, we will calculate the average RSS value from each AP. If the value is below the threshold we previously set, the corresponding AP is judged shut off, and we will replace the AP with another AP in the alternative list. The AP list will be used in the next updating.

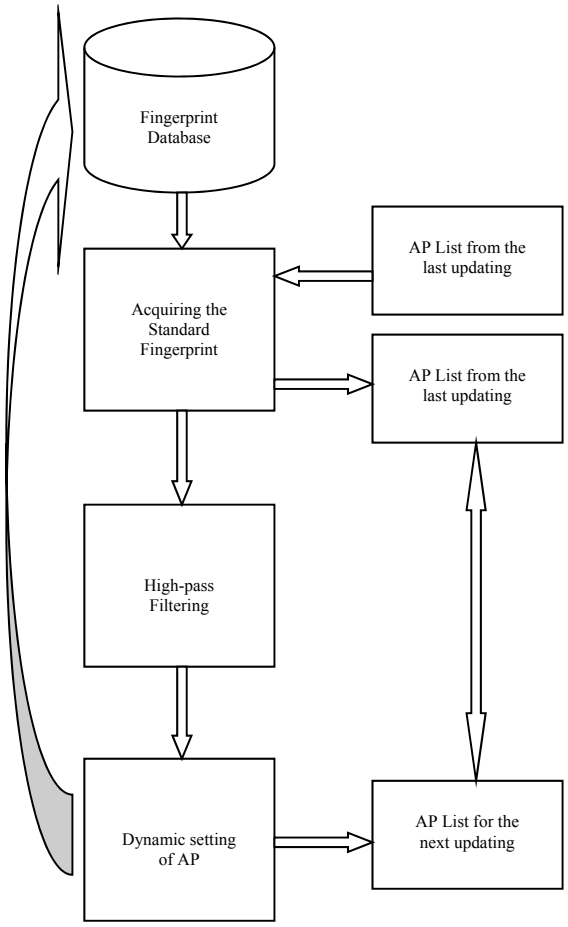
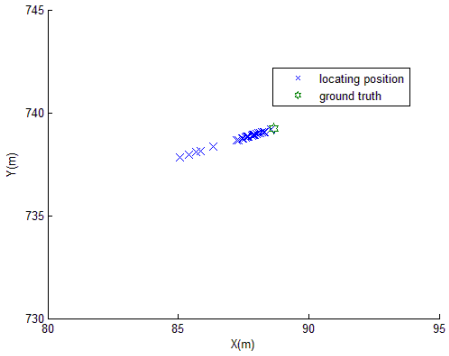
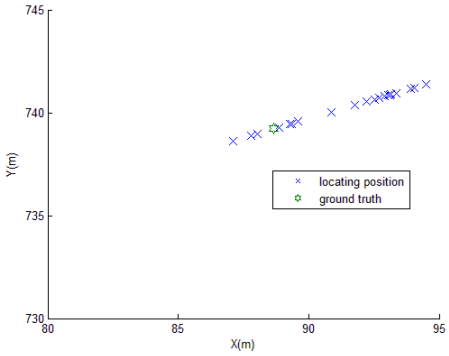


Fig. 13. Flow diagram of database updating based on Wi-Fi fingerprints' reliability model

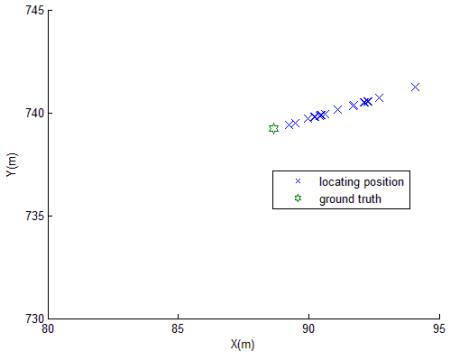
The flow diagram of the updating based on fingerprints' reliability model is shown in Figure 13. Figure 14 shows the positioning performance of updating based on Wi-Fi fingerprints' reliability model when the fingerprints in the updating data are less than the ones in the original database. Before the first updating, we shut off two of the original APs. So in Figure 14 (b), we can see that the positioning accuracy deteriorate a little. However, after another two times of updating, the positioning performance recovers as is shown in Figure 14 (c) and (d).



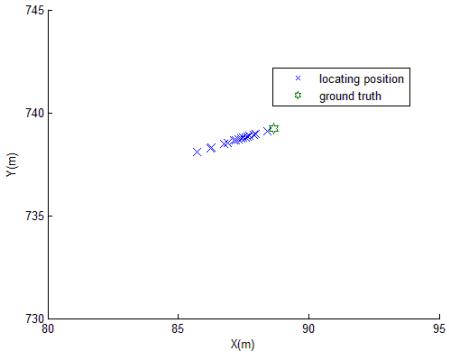
(a)



(b)



(c)



(d)

Fig. 14. Positioning performance of updating based on Wi-Fi fingerprints' reliability model (a) Initial database (b) Database of the first updating (c) Database of the second updating (d) Database of the third updating

IV. EXPERIMENTAL RESULTS

To validate the performance of our proposed method, several experiments in different Wi-Fi environments were conducted. The experiment sites were situated at our laboratory building. In the experiment, we set the updating interval to 12 hours. We updated the fingerprint database by using the crowdsensing data from the users in the coverage area every 12 hours, then we did real-time positioning with the new database.

A. Stable Wi-Fi environment

In stable Wi-Fi environment, all the APs are normally running and the distribution of indoor obstructions is fixed. To inspect the robustness of the database updating, we set the amount of the fingerprints in the updating data smaller than the one in the original database, M/K is 3 in (20). The reliability model is:

$$\mathrm{R(t)} = \begin{cases} e^{-3.725\times10^{-8}(t_0-t)} & t_0 - t < 15days \\ 0 & t_0 - t \geq 15days \end{cases} \tag{21}$$

After the original database was generated, we did 8 times of database updating with the two methods above. First, we calculated the relative updating ratio and the absolute updating ratio. The relative updating ratio is the ratio of the fingerprints in the updating data which are elected as fingerprints in the new database. The absolute updating ratio is the ratio of the fingerprints in the initial database which are eliminated.

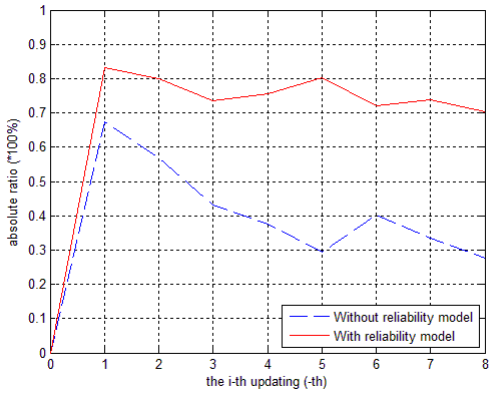


Fig. 15. Comparison of relative updating ratio in stable Wi-Fi environment

In Figure 15 and 16, the peaks of relative updating ratios both appear in the first updating, because they both corrected the initial database a lot in the first updating. After the first updating, both curves became flat. It's easy to find that the method with reliability model performs better in updating efficiency.

We measured the average positioning errors of the location systems with the two database updating methods and without updating. Figure 17 shows the result. In stable Wi-Fi environment, database updating based on the reliability model has a similar influence on positioning accuracy to the updating without reliability model. They both can improve the positioning accuracy with several times of updating. The average positioning error with the initial database is 4.10 m.

After 8 times of updating, updating without the reliability model reduced it to 1.70 m and updating with the reliability model reduced it to 1.88 m. Figure 18 shows the position variance comparison. Both the methods can converge the positioning results by the standard fingerprints and high-pass filtering based on fingerprints' similarity. After 8 times of updating, the positioning variances were reduced to 3.65 and 3.22 by the 2 methods.

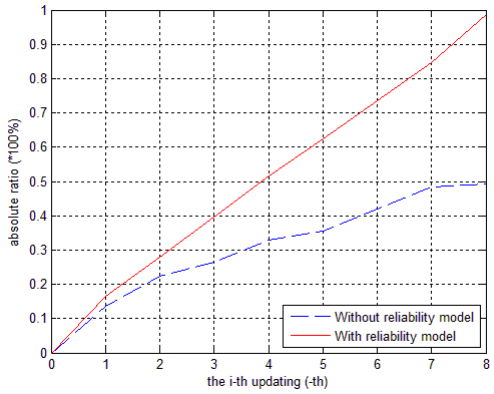


Fig. 16. Comparison of absolute updating ratio in stable Wi-Fi environment

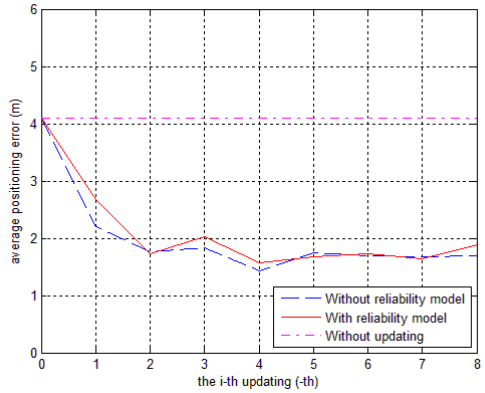


Fig. 17. Comparison of average positioning error in stable Wi-Fi environment

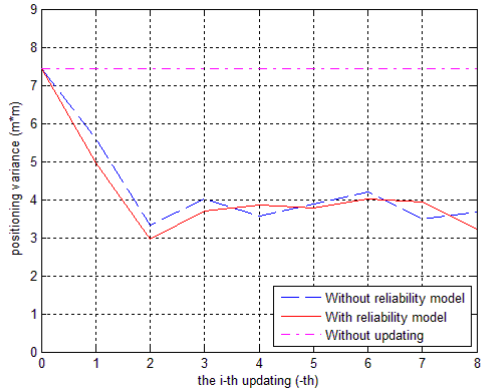


Fig. 18. Comparison of positioning variance in stable Wi-Fi environment

B. *Unstable Wi-Fi environment*

In unstable Wi-Fi environment, the Wi-Fi signal strengths from the APs may change owing to the replacements of APs or the changes of indoor obstructions. In our experiments, we did 10 times of updating and we shut off 30% of the APs before the 3rd , 6th, 9th updating to simulate the change of the Wi-Fi environment and inspected the performances. Other parameters are the same as the experiment in stable environment.

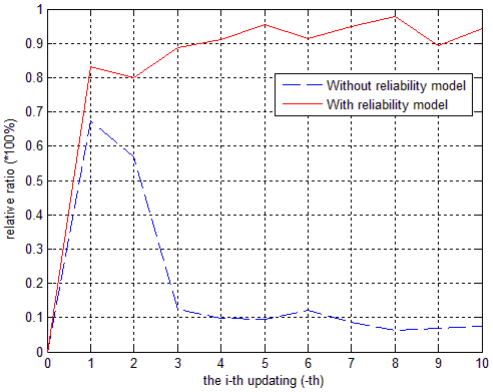


Fig. 19. Comparison of relative updating ratio in unstable Wi-Fi environment

Figure 19 and 20 show the relative and absolute updating ratios of the two methods. From Figure 19, we find that the relative ratio of the database updating based on fingerprints' similarity reduced sharply when the Wi-Fi environment changed. On the contrary, the relative ratio of database updating based on the reliability model increased when the change happened.

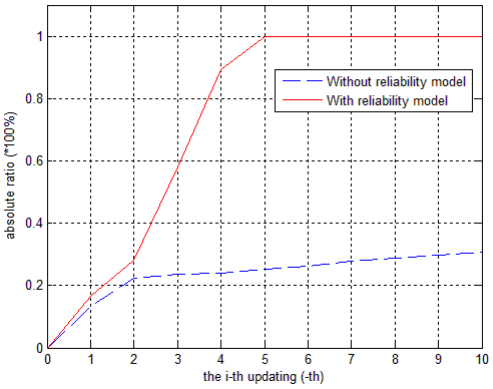


Fig. 20. Comparison of absolute updating ratio in unstable Wi-Fi environment

We measured the average positioning errors and the positioning variance of the location systems with the two database updating methods and without updating. Figure 21 and 22 show the result. When the change of the Wi-Fi environment happened, the database updating based on fingerprints' similarity couldn't update the fingerprint database successfully. The different fingerprints brought by the change of the Wi-Fi environment are filtered out by mistake. The average positioning error and positioning variance increased as we shut off more and more APs. When

there were only 10% APs remained, the average positioning error was 5.87 m and the positioning variance was 13.25. On the contrary, the database updating based on the reliability model maintained the database successfully when the change of the Wi-Fi environment happened. The positioning performance was influenced right after the change happened, but 2 or 3 times of updating reduced the positioning error and the variance. From the curves in Figure 21 and 22, we find that the positioning performance with updating based on the reliability model remained stable. Even though 90% of the original APs were replaced, the reliability model and the dynamic AP setting help maintain the database and make the fingerprint database organic. After 10 times of updating, the average positioning error was 2.14m and the variance is 3.57.

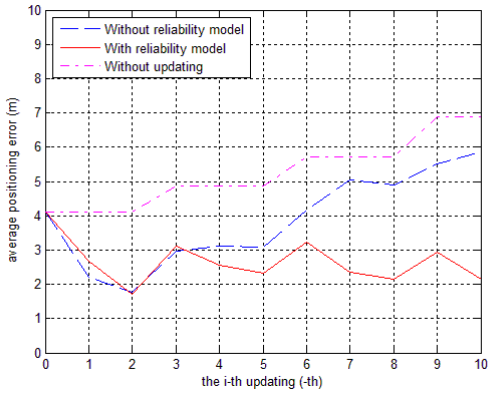


Fig. 21. Comparison of average positioning error in unstable Wi-Fi environment

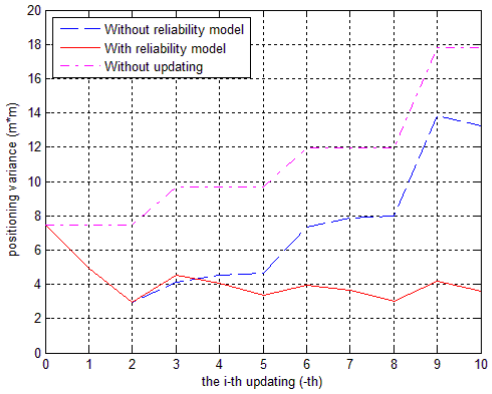


Fig. 22. Comparison of positioning variance in unstable Wi-Fi environment

V. CONCLUSIONS

We present two fingerprint database updating methods: Updating based on fingerprints' similarity and Updating based on the reliability model. The experimental results demonstrate that both the methods can be utilized to maintain the fingerprint database in stable Wi-Fi environments. The updating based on the reliability model has a more robust performance in unstable Wi-Fi environments and it can make the fingerprint database organic with the utilization of crowdsensing data. In future, the algorithm will be more intelligent with the cooperation of more advanced signal sampling and users' feedback.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China under Grant 61573242 and 61402283, partly funded by the Shanghai Science and Technology Committee under Grant 15511105100 and 16DZ1100402, and the National Science and Technology Major Project under Grant GFZX0301010708.

REFERENCES

[1] Pei, L., Chen, R., Liu, J., Kuusniemi, H., Tenhunen, T., & Chen, Y. (2010). Using inquiry-based Bluetooth RSSI probability distributions for indoor positioning. Journal of Global Positioning Systems, 9(2), 122-130.

[2] Chen, L., Pei, L., Kuusniemi, H., Chen, Y., Kr02ger, T., & Chen, R. (2013). Bayesian fusion for indoor positioning using bluetooth fingerprints.Wireless Personal Communications, 70(4), 1735-1745.

[3] Pei, L., Chen, L., Guinness, R., Liu, J., Kuusniemi, H., & Chen, Y., et al. (2013). Sound positioning using a small-scale linear microphone array.Indoor Positioning and Indoor Navigation (IPIN), 2013 International Conference on . IEEE.

[4] Pei, L., Chen, R., Liu, J., Tenhunen, T., Kuusniemi, H., & Chen, Y. (2010). An Inquiry-based Bluetooth indoor positioning approach for the Finnish pavilion at Shanghai World Expo 2010. Position Location and Navigation Symposium (PLANS), 2010 IEEE/ION (Vol.298, pp.1002 - 1009). IEEE.

[5] Chen, R., Kuusniemi, H., Hyyppae, J., Zhang, J., Takala, J., & Kuittinen, R., et al. (2010). Going 3d: personal nav and lbs. Gps World, 21.

[6] Qian, J., Ma, J., Ying, R., Liu, P., & Pei, L. (2013). An improved indoor localization method using smartphone inertial sensors. Indoor Positioning and Indoor Navigation (IPIN), 2013 International Conference on (pp.1 - 7). IEEE.

[7] Qian, J., Pei, L., Zou, D., Qian, K., & Liu, P. (2014). Optical flow based step length estimation for indoor pedestrian navigation on a smartphone.Position, Location and Navigation Symposium - PLANS 2014, 2014 IEEE/ION (pp.205 - 211). IEEE.

[8] Chen, R., Pei, L., & Chen, Y. (2011, September). A smart phone based PDR solution for indoor navigation. In Proceedings of the 24th International Technical Meeting of the Satellite Division of the Institute of Navigation (pp. 1404-1408).

[9] Qian, J., Pei, L., Ying, R., Chen, X., Zou, D., Liu, P., & Yu, W. (2014). Continuous motion recognition for natural pedestrian dead reckoning using smartphone sensors. In Proceedings of the 27th International Technical Meeting of the ION Satellite Division, ION GNSS.

[10] Liu Qianchen, "Research on WLAN-Based Indoor Fingerprint Location Technology," Master dissertation, Shanghai Jiao Tong Univ., Shanghai, P.R.China, 2013.

[11] Yuan, Y., Pei, L., Xu, C., Liu, Q., & Gu, T. (2014, November). Efficient WiFi fingerprint training using semi-supervised learning. In Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), 2014 (pp. 148-155). IEEE.

Thomas, G, Binghao, L, Andrew, G, & Chris, R. (2010). Database updating through user feedback in fingerprint-based Wi-Fi location systems. In Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), 2010 (pp. 1-8). IEEE.