Some optimizations of WiFi-Based Indoor Positioning

Hanqing Liu,Shanghai JiaoTong University

Recently, several indoor positioning methods have been proposed. Due to the high cost and complexity of some solutions based on Bluetooth, Infrared and Ultra-sound, even though they are able to achieve a higher accuracy, they are not widely adopted and only exist in theory or some specific situation. Meanwhile, indoor positioning based on WiFi has caused great attention because it make used of the current access point instead of deploy new ones like infrared or ultra-sound. This paper gives a survey of two popular indoor positioning methods based on WiFi received signal strength and fingerprint which uses received signal strength to generate an unique fingerprints. In our work, we study and compare these two different approaches and the optimizations of them and then we present our work.

***Index Terms*—Indoor positioning, WiFi-based, SLAM,Received signal strength .**

I. INTRODUCTION

INDOOR positioning system(IPS) has attracted much attention in the past decade because of its social and commercial values. An indoor positioning system enables a mobile device to get its position in an indoor environment. The most common positioning tool is global positioning system (GPS). However, it cannot be used in indoor environment, because it requires light-of-sight transmission between user and satellites which is not available indoor. At the meantime, the indoor environment is far more complex than the outdoor one. The existence of obstacles such as walls and furniture or even human beings makes it difficult to locate a person precisely. Thus, compared with the global positioning system (GPS) which is the most widely use of outdoor positioning, indoor positioning has the feature of complex, non-line-of-sight, more complicated obstacles, signal fluctuation or noise, environmental changes.

Due to the high demand of the indoor positioning system and the feature environment, several approaches have been discovered. Some articles [1], [2] have introduced an overview of various technology options which have been studied or put into practical use for the design of an indoor positioning system such as radio-frequency identification (RFID), infrared (IR), wireless local area network (WLAN), ultrasound, Bluetooth, vision analysis and audible sound. These approaches have some features which can highly fit a specific environment. Due to the advantages and limitation of these approaches, different indoor positioning have been developed based on one technology or a combination of several technologies above to eliminate the limitation of the system.

Among all the approaches, some of the technologies like IR, ultrasound has a fatal limitation because the deployment of the system based on them will be expensive even though compared with other approached based on WiFi or RFID, the accuracy of IR and ultrasound is much higher because of the adoption of time-of-arrival or angle-of-arrival scheme. If the accuracy is not highly demanded by users which means users only want to have a vague idea of their location in the room, wifi based inddor positioning is the best approach to indoor localization in this case.

The remainder of this paper is organized as follows. The related work of wifi based indoor positioning is presented in section 2. Section 3 will give an overview of current optimizations of indoor positioning based on received signal strength and fingerprint. In section 4, we will describe our implement of indoor positioning based on received signal strength only and explain the reason why the fingerprint is not adopted in our approach. I will also show that what should be noticed in a practical case. Finally, section 5 will draw the conclusion of our work and give a future prospect of indoor positioning system.

II. RELATED WORK

In this section, we first introduce the basic principle of wifi-based indoor positioning and the LDPL model which the RSS only indoor positioning system is based on (section 2-A). Then we will describe the basic structure of fingerprint-based indoor positioning system (section 2-B). Other related work and the similarities and differences of the two wifi based system will be presented in the third part (section 2-C).

A. RSS based IPS and LDPL model

A LDPL model can be used to predict RSS at various locations in the indoor environment [6]. It can also be used to predict the distance between the user and the access point(AP). Many study have shown that the signal indoor path loss follows the LDPL model.

$$P_{ij} = P_i - 10\gamma log d_{ij} + R \tag{1}$$

where p_{ij} represents the signals path loss when the distance between the j^{th} user and i^{th} AP, p_i represents the distance is one unit (1m) to i^{th} AP, R is the path loss parameter which depends on the environment. R is a random variable which follows normal distribution. LDPL equation (Equation 1) is a system of simultaneous non-linear equations. From the equation, we can clearly draw the conclusion that if p_i and γ is known, then p_{ij} can be converted to the distance d_{ij}. If these two parameters are certain, we only need 3 AP to locate the user.However, in real case, it is difficult to determine every parameter for each AP because the environment is keep changing indoor, so we need at least five AP to locate the user so that the demand for accuracy can be met in some aspect.

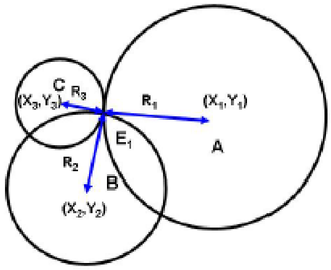


Fig. 1. Triangulation positioning.

After determining the value of d_{ij} respectively from the value of p_{ij}, the approach to locating the user is almost done. The basic of positioning algorithm based on RSS only is triangulation position technique. The principle of triangulation technique is shown in Figure 1. If the location of the 3 APS is known, the users position can be determined by the length of R1, R2, R3. With this method, another problem should be noticed. The received signal strength from a certain AP at a certain point is not always a fixed value. It usually follows normal distribution. Sometimes it will cause huge error of the positioning. Thus, a series of work should be done to eliminate the error. We will present our work in section 4.

B. Basic structure of fingerprint-based indoor positioning system

Fingerprint-based indoor positioning system is first introduced by Bahl in 2000. It is used to improve the accuracy of indoor positioning. It can achieve a relatively small error which positioning based on RSS only cannot achieve. As a result, a large part of indoor localization system adopts fingerprinting as the basic of their algorithm. The main idea is to use pre-measuerd data. Fingerprint-based positioning system includes two phases: offline training phase and online position determination phase [9]. Due to the multi-path effect, user will receive a unique signal because of the combination of signals from different paths. The received signal is unique in radio frequency response and signal strength. The unique frequency response is usually used as the fingerprint. In offline training phase, the fingerprint of every location is collected and stored into the database. In online position determination phase, when user receive a signal, it determines its position by searching the fingerprint in database.

Even though we have mentioned that fingerprint-based positioning is more precise that RSS-based positioning, it is not always perfect. According to Hongbo Liu[11], large error may occur because two roots: permanent indoor environment factors such as walls, furniture and obstacle which can prevent the radio signals propagation and measurement mismatch between testing and training data. These error motivates some advanced schemes which will be introduced in section 3.

III. An overview of current optimizations of Wi-Fi Fingerprint-Based Indoor Positioning

The presence of fingerprint-based indoor positioning is a large step forward in the field of indoor positioning. However,

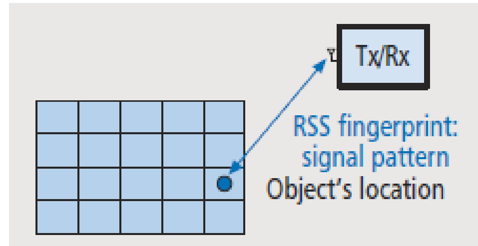


Fig. 2. Fingerprint-based positioning with grid.

as shown in section 2-B, it does not mean that the scheme of indoor positioning is perfect.

Besides the problem of large errors, the pre-work to build database makes the indoor positioning much too complex. As shown in Figure 2, the fingerprint of each grid must be measured in advance. To achieve a certain accuracy, the grid should be small enough. For example, if the desired accuracy is less than 2 meters, the diameter should at least less than or equal to 2 meters. In practical case, the diameter should far less than 2 meters. Thus, if the indoor area is large, a lot of pre-work should be done in advance. Besides, the components of fingerprint make the case even worse. Based on P.Bahl's design in 2000[10], each fingerprint contains the signal strength from each base station or AP, if the number or location of the APS changes while the database is not updated in time, the location result will be a huge mistake. Meanwhile, the update of database is as complicated as the measurement at the first time even only a small part of the APs have changed.

The current optimizations are basically from these two aspects: The complexity of pre-mesurement and the error caused by indoor environment and mismatch. We will present two optimizations SLAM and peer assist positioning in section 3-A and 3-B.

A. Simultaneous Localization and Mapping (SLAM)

Simultaneous localization and mapping is a technique which can locate the user and collect the data and update it to the database at the same time. Actually, SLAM is a technique which is first used in the field of computer vision. In the field of positioning, the concept of SLAM is a little different. WiFi-SLAM uses the Gaussian process latent variable models to relate RSS fingerprints and models human movements[8]. In current work[8], locations are computed through the deterministic MDS method.

Figure 3 shows the basic system structure of the SLAM system. It also has two phase: the training phase and the operating phase. However, the training phase is not a pre-work which is different from traditional fingerprint based localization. When the indoor environment is first visited, the environment is unknown to the database. Then the system is in the training phase. In this phase, user cannot get his location information because the database is almost empty. When the user is walking around, the raw received signal strength is collected on his path. Through multidimensional scaling, the system generates a fingerprint space. If the user has walked long enough, he will provide several row fingerprint data.

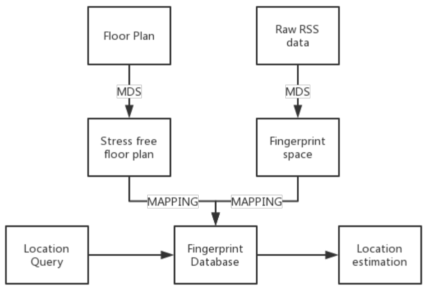


Fig. 3. Structure of SLAM system.

Through MDS, the similarity of the data would be found and their intersection can be inferred. At the mean time, the floor plan is generated by different kinds of sensors in users mobile device so that combining the result of fingerprint space, the route can be constructed together to generate the construction of the environment. The construction will be modified if users keep walking around until the whole map of the indoor environment is generated.

A floor plan shows a view of a building structure from above, including the relationships between rooms, spaces, and other physical features. The geographical distance between two locations in a floor plan is not necessary to be the walking distance between them due to the block of walls and other obstacles.

RSS fingerprints are collected during users' routine indoor movements. Users walk in a building and their mobile phones record RSS fingerprints along their walking paths, as well as the footsteps between every pairs of two consecutive fingerprints. After we collect the fingerprints $F = f_i, i = 1...n$ and the distance matrix $D = d'_{ij}$, some pre-processing should be done because the movements of human beings are irregular and the data pre-processing helps to merge similar fingerprints. Usually, given two fingerprints f_i and f_j, we use following equations to define their RSS similarity:

$$\phi_{ij} = ||f_i - f_j||_i = \sum_{k=1}^{m} |s_k - t_k| \tag{2}$$

For f_i and f_j , if their dissimilarity ϕ_{ij} is smaller than a predefined threshold Th, then they are merged as a same point in the fingerprint space to be generated. Otherwise, if $\phi_{ij} > Th$,$f_i and f_j$ are treated as two different points. The determination of epsilon is based on the fingerprint samples collected at a given location .

In the operating phase, the users can finally get their position because the database has been updated by themselves. It is an efficient method but trades of the accuracy in that the construction of the map of indoor environment is not precise.

B. Peer Assisted Localization

The peer assisted localization is another kind of optimization of fingerprint- based localization. Due to the two reasons we have explained above: the complexity of indoor environment and the mismatch, the position of user generated by fingerprint is not necessarily the real location of the user. It poses a new problem: How can the system know that the position of user is wrong. The peer assisted localization [11] enables the system to realize the existence of error by introducing a scheme which users help each other.

The realization of peer assisted localization is based on the assumption that in the indoor environment, there are a lot of users who are willing to help each other so that they can improve their own positioning accuracy as well. This assumption is reasonable in many situations. The assistance provided by peers is implemented by time of arrival scheme based on sound signals.

The work flow of the system will be described as follows. When a user needs to increase the accuracy of his estimated location, the device will broadcast a special audio signal. When the peer devices which also want to improve the accuracy accept the signal, they will send an audio signal back immediately. They will repeat send the signal until the distance between each other is known by the server. The server collects all the devices data and calculates the distance between the devices. Based on the distances, the server computers the location estimation of all the devices and sends back the result.

Generally, the algorithm has two parts: generate the graph based on the distance between the devices and rotate or translationally moves the graph on the signature map. In the processing of generate the graph, we follows the rigid graph theory: A complete graph with a distance between any two vertices is rigid. This theory implies that if n devices want to be located precisely, the time complexity of the generation of graph is $O(n^2)$.

a) Step1: Rotation: In the second part, the rotation contains 3 part: Compute edge directions from acoustic ranging; Compute edge directions from initial wifi localization; Estimate the orientation of the graph.

Suppose we take the L longest edges to avoid a large error and we get 2 L-dimension vectors. Denote the two vector v_l and v_l' where v_l means the edge direction of the initial wifi localization and v_l' means the edge direction of the acoustic ranging localization. The server will computer a Φ which can minimize the value of inner product summation between v_l and v_l'.

$$argmax \sum_{l=1}^{L} v_l v_l'^T \tag{3}$$

Due to ranging errors, the real orientation may differ from Φ.To ensure that the true orientation is covered, the search of Φ is conducted in an orientation range of [Φ - Δ Φ, Φ + Δ Φ].

b) Step2:Translationally Moves: After the rotation, the server tries to impose the graph G onto the WiFi signature map so that each vertex is restricted into a small grid as shown in Figure 2. Denote pi the initial localization of the devices and qi the new estimated localization of the devices. With a movement of alpha meters, the sever search for the optimal location which fits Equation 3.

$$argmin \sum_{i=0}^{M} [f(q_i) - f(p_i)][f(q_i) - f(p_i)]^T \quad (4)$$

The peer assisted localization can significantly improve the accuracy of the indoor positioning by decreasing the error by up tp 50 percent according to Hongbo Liu who proposes the algorithm [11]. However, he also notices some conflicts in two goals of indoor positioning : bing more accurate and decreasing the complexity of the algorithm. The trade of between the two goals will be further explained in the last section. In peer assisted localization, there are a few demands which can be the limitation of the algorithm. a) The time complexity of the algorithm is high since every edges between each pair of vertexes should be calculated so that there exists requirements to the server to ensure the calculation. b) The peer involvement. Without a number of peer involvements, the accuracy cannot be guaranteed so that this indoor positioning system is not suitable for the case where users are not willing to publish their own position and help each other. c) The movement of user. If the user is keep moving, then every time the users want to get his own position, the server must regenerate the graph since most of the edges of the graph has changed. The frequent movement of users gives the server a heavy load.

IV. Our implement of indoor positioning based on RSS only

Our implement of indoor positioning is based on a project which cooperates with FuShikang. The implement of this kind of practical project is different from the research in the lab since the feasibility, cost and the practical environment should also be considered while in lab, the cost, environment and other conditions can all be some assumptions. The difference in the essence of lab work and a practical project leads researchers different attitude towards the research. For instance, in our project, we have considered some advanced technique such as infrared and ultra-sound to enhance the accuracy however due to the deployment is expensive and the gain from positioning is not so considerable.

At the meantime, due to the widely deployment of wifi routers in the area, the wifi-based technique comes into our mind. Then the remaining problem is to decide whether to use RSS based positioning or fingerprint-based positioning. The advantage of fingerprint-based positioning over RSS based positioning is clear: accuracy. However, in the practical job, weve noticed that the environment in which the indoor positioning system will be deployed is a factory and the area is quite large. The large area is a great challenge to the offline phase since the area can be divided into thousands of grid and in each grid a series of data should be collected in advance and due to the router can be moved or added, the database should be updated frequently. So considering the fact that positioning based on fingerprint may cause a lot of unnecessary work, we decide to use the tranditional method: indoor positioning based on RSS only and use the combination of triangular positioning and other optimizing methods.

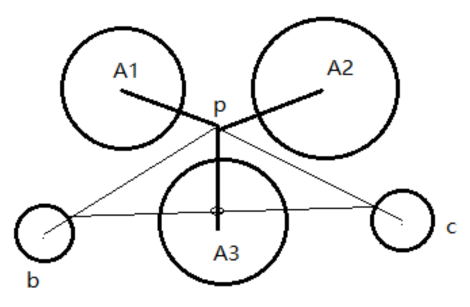


Fig. 4. Illustration of our positioning algorithm.

A. *The algorithm of our indoor psotioning system*

The main idea of our positioning algorithm is based on Figure 1 which is the basis of position. We calculate the distance between user and the anchor according to the LPDL model. The selection of triangular positioning is to acquire the users position as soon as possible since the number of users in a factory is nearly a thousand. The primary goal of the system is to locate the user as fast as possible with a relatively low demand in accuracy. However, positioning based on 3 point is far too risky. The signal strength received by user performs as a normal distribution. A large error which is not affordable can occur at any time. Thus some optimization is definitely necessary.

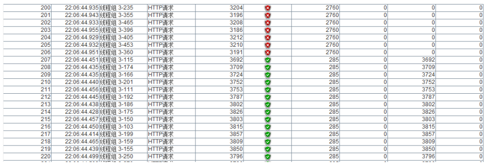
1) *Step 1: Triangular positioning*

At first, the user gets the data of signal strength from$AP_1, AP_2...AP_n$ and forms a vector $RSS_1, RSS_2...RSS_n$. In this vector, we take 3 random APs whose value of signal strength is greater than or equal to a certain threshold $Th_1$and denote them as $RSS_i, RSS_j andRSS_k$ where their location is denoted by $L_i, L_j andL_k$. With the value of $AP_i, AP_j andAP_k$, the distance $d_i, d_j andd_k$ is calculated based on the LDPL model. A factor should be noticed is that different from the ideal case shown in Figure 1, the cross point may not occur or many cross points may occur. In most of the case, the cross point may not occur so we need to find a point where the user has the most chance to locate in. The sum of the distance between the user and the center of the circles i, j and k will help. When the sum has a minimum at a point P_i, we can take the pointP_i as the cross point in practice.

2) *Step 2: Optimization*

Due to the large error caused by the instability of the received signal strength, some optimization should be applied to avoid the error. Consider the following case: At time t_0, the user was located to p_0 at the south-east corner of the area and at t_1, the user sends his received signal strength to the server again. At this moment, the server surprisingly finds that the users location has changed to the north-west corner of the area based on the selected AP value and the area is so large that the user cannot move such a long distance in the time interval $t_1 - t_0$. A mechanism should be made to ensure the server to be aware of the mistake.

Thus, during the process of locating, some schemes should be applied. In our, algorithm, for each vector



200	22:06:44.935	线程组 3-235	HTTP请求	3204		2760	0	0	0
201	22:06:44.943	线程组 3-355	HTTP请求	3196		2760	0	0	0
202	22:06:44.933	线程组 3-465	HTTP请求	3208		2760	0	0	0
203	22:06:44.955	线程组 3-396	HTTP请求	3186		2760	0	0	0
204	22:06:44.929	线程组 3-405	HTTP请求	3212		2760	0	0	0
205	22:06:44.932	线程组 3-453	HTTP请求	3210		2760	0	0	0
206	22:06:44.951	线程组 3-360	HTTP请求	3191		2760	0	0	0
207	22:06:44.451	线程组 3-115	HTTP请求	3692		285	0	3692	0
208	22:06:44.435	线程组 3-174	HTTP请求	3709		285	0	3709	0
209	22:06:44.435	线程组 3-166	HTTP请求	3724		285	0	3724	0
210	22:06:44.440	线程组 3-201	HTTP请求	3752		285	0	3752	0
211	22:06:44.456	线程组 3-111	HTTP请求	3753		285	0	3753	0
212	22:06:44.445	线程组 3-192	HTTP请求	3787		285	0	3787	0
213	22:06:44.438	线程组 3-186	HTTP请求	3802		285	0	3802	0
214	22:06:44.428	线程组 3-175	HTTP请求	3826		285	0	3826	0
215	22:06:44.457	线程组 3-150	HTTP请求	3803		285	0	3803	0
216	22:06:44.450	线程组 3-103	HTTP请求	3815		285	0	3815	0
217	22:06:44.414	线程组 3-199	HTTP请求	3857		285	0	3857	0
218	22:06:44.465	线程组 3-159	HTTP请求	3809		285	0	3809	0
219	22:06:44.439	线程组 3-155	HTTP请求	3850		285	0	3850	0
220	22:06:44.499	线程组 3-250	HTTP请求	3796		285	0	3796	0

Fig. 5. A running stress test which returns some errors

$RSS_1, RSS_2, , RSS_n$, we choose the greatest value among n received signal strength. We denote it RSS_{max} and compared it with another threshold $Th_2 and Th_3 (Th_2 > Th_3 > Th_1)$. If RSS_{max} is greater than Th_2 which is a rare case, then the distance between the AP with RSS_{max} and user is so close that they are almost at the same location. Then, the position of user can be equivalent to the position of the AP. If RSS_{max} is greater than Th_3 and less than Th_2, it means that the location of the user has some relationship with the AP with RSS_{max}. At this time, we take the middle of the estimated location and the location of the AP as the real location of user. If the threshold is greater than TH1 but less than Th_3, it means that the AP has a little thing to do with the location of the user, in this time, we calculate the location based on the estimated location, the APs location and the users location last time the server recorded.

B. *Some follow-up work of our project*

After the algorithm is done, we can test it in small scale to make sure its performance. However, the large-scale deployment is still not available since there is some other follow-up work which has not finished yet. We have mentioned above for a series of time that our work is a practical work so that we have to consider the real performance of every part of the system but not put our emphasis on theoretical assumption. For instance, the structure of the database and the performance of the server are both problems we need to concern about.

An important part of job is the stress test of the server by which we can know the limitation of the server and assure the range of load impressed onto the server in the real case.In the processing of press test, the only available data is the response of the server and the error log and access log saved in the server. The data helps us to analyze the fault in the algorithm or the interface. In our project, we use Jmeter to simulate the user behaviors by creating hundreds of or thousands of threads which can post request to the server at a certain speed in a fixed time interval. The data we post is not random, it should be processed by the server so that the test can receive a result which is close to reality.

Since the http request is post to the server through TCP protocol, the bandwidth can sometimes be the bottleneck of the test which should be dealt with carefully. Luckily, in our test, the bottleneck is not our bandwidth. In our early test, we found some legal problems which strictly restrain the performance of the server even if the usage of the CPU of the server is not an unacceptable number. Through many test, we have ensured that the bottleneck of our system is the entrance of server. The frequently access to the database may sometimes cause the read and write problem which leads to the congestion even if the database supports multi-thread entrance. When several threads want to enter the database and change the value of the same object, the read and write problem may occur and the threads must wait in a line. It is similar to the case that the database only supports a single thread. When the line is long enough and occupies all the buffer, the error occurs. In our test, because of the limitation of the entrance of server, the post rate is limited to about 100 times per second which is far from enough in the real case. We are still working on this problems by understanding the essence of the queuing problem and optimize the structure of database so that the read and write problem will not happen so frequently.

V. Conclusion and Future work

Indoor positioning is a perspective field in the future because of the widespread of WiFi access points and mobile device. In addition, with the increasing of the complexity of the indoor environment, this technique will be more and more widely used. Meanwhile, the demand for the accuracy and complexity of the algorithm will certainly be increasing. In this paper, we discussed WiFi-based indoor positioning system based on RSS only and fingerprint scheme and analyzed the advantage and disadvantage of both of the technique.

In section 4, we introduced our work based on RSS only and explained the body of our algorithm. Since in the practical work, the performance of server and the restriction of the area, the server and the environment should be considered, the algorithm we adopted is not too complicated but it performed well in our test. In the future work, we may eliminate the limitation of the database and further optimize our system.

Acknowledgment

I would like to thank Prof. Wang and Prof. Tian for their guidance in this semester. Prof. Wang and Prof. Tian showed our projects direction at the beginning the semester and offered us some necessary equipment in their laboratory. Meanwhile, I also want to thank Yang Li and Tiance Wu who work in Prof. Tian's laboratory. Their assistance is identically important. The main frame and the idea of the project is from their work with which they lead us into the area of indoor positioning.

References

[1] M. Vossiek, L. Wiebking, P. Gulden, J. Wiehardt, C. Hoffmann and P. Heide, *Wireless Local Positioning*. IEEE Microwave Mag,vol 4,Issue 4: December 2003.

[2] J. Hightower and G. Borriello, *Location sensing techniques*. Technical Report UW CSE, Department of Computer Science and Engineering, University of Washington, 2001.

[3] Yanying Gu and Anthony Lo, *A Survey of Indoor Positioning Systems for Wireless Personal Networks*. In IEEE COMMUNICATIONS SURVEYS AND TUTORIALS: 2009.

[4] Chouchang Yang and Huai-Rong Shao, *WiFi-Based Indoor Positioning*. In IEEE Communications Magazine: 2005.

[5] Suining He and S.-H. Gary Chan, *Wi-Fi Fingerprint-Based Indoor Positioning:Recent Advances and Comparisons*. In IEEE COMMUNICATIONS SURVEYS AND TUTORIALS: 2016.

[6] Krishna Chintalapudi, Anand Padmanabha Iyer and Venkata N. Padmanabhan, *Indoor Localization Without the Pain*. In MobiCom: 2010.

[7] Y. Gwon and R. Jain, *Error Characteristics and Calibration-Free Techniques for Wireless LAN-based Location Estimation* . In Mobiwac: 2004.

[8] Zheng Yang, Chenshu Wu and Yunhao Liu, *Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention* . In MobiCom: 2012.
[9] K. Kaemarungsi and P. Krishnamurthy, *Properties of indoor received signal strength for WLAN location fingerprinting* . Proc. 1st Annual International Conference on Mobile and Ubiquitous Systems: 2004.
[10] P. Bahl and V. N. Padmanabhan, *RADAR: An in-building RF-based user location and tracking system* . In Proceedings of the IEEE International Conference on Computer Communications (INFOCOM'00): 2000.
[11] Hongbo Liu, Yu Gan and Jie Yang, *Push the Limit of WiFi based Localization for Smartphones* . Push the Limit of WiFi based Localization for Smartphones: 2012.