Indoor 2.5D Positioning of WiFi Based on SVM

Shuai Zhang*, Jiming Guo, Wei Wang, Jiyuan Hu
Research Center for High Accuracy Location Awareness
School of Geodesy and Geomatics
Wuhan University
430079, Wuhan, China
*Corresponding author: 2016102140015@whu.edu.cn

***Abstract*—With the rapid development and popularization of WiFi technology, indoor positioning technology based on WiFi has become a hot spot. At present, research on WiFi is mainly focused on two-dimensional positioning in the field of indoor positioning, which is less concerned about the 3D positioning. In this paper, a method of WiFi indoor 2.5D positioning based on support vector machine (SVM) is proposed. The method first uses the WiFi signal of each floor to identify the floor by SVM, and which obtains the position along the altitude direction. Then, the weighted k-nearest neighbor (WKNN) algorithm is applied for plane location estimation over a set of WiFi location fingerprints. Thus, indoor three-dimensional positioning is realized. Through improvement of the AP selection method, the accuracy of plane positioning is enhanced. The experimental results show that the recognition accuracy rate of the floor is 99.09%, and the average error of the indoor location estimation is 0.63m.**

Keywords*—WiFi fingerprint ; AP selection ; floor recognition；support vector machine；indoor positioning

I. INTRODUCTION

Application demand of location-based services (LBS) is increasing rapidly, and LBS has been integrated into our daily life and learning. At present, global navigation satellite systems (GNSS) have provided sub-meter level accuracy in many applications, which have a good performance in outdoor positioning. However, the accuracy of GPS can't meet the need of the people due to signal fading and multipath effect in the indoor environment [1-3]. People are more and more interested in indoor positioning these years. Because indoor positioning based on WiFi does not require additional equipment, which has become a research hot spot [4].

At present, a variety of complex structures come into being along with the gradual improvement of people's living standards [5]. People can't easily find their location in a complex indoor environment such as large shopping malls, exhibition hall, etc.

In order to get the position along the altitude direction and provide better location service, people have already done some research on floor recognition. Li et al. [6] presented a method of floor identification of the building, which based on WiFi received signal strength indication (RSSI). The method first records all access points (AP) belonging to floor. By statistics of the number of AP greater than the threshold of RSSI in each floor, if the number of AP is the most in one layer, the floor is defined as the current floor. The location method reduces computation, but the location result will be affected if marked APs reduce or move to other floors. Additionally, WiFi fingerprints can also be used to identify the floors. This way of floor recognition based on WiFi fingerprints is to scan all the fingerprint points and calculate the Euclidean distance from test points to all fingerprint points. The floor of the nearest fingerprint point at the distance test point is current floor, but the method of WiFi fingerprint to identify the floors is a lot of time-consuming. In order to resolve this problem, Deng et al. [7] proposed a method based on K-means algorithm for identification of the floor. This method uses spatial clustering algorithm, and greatly reduces the computation time. But this method needs to determine the number of clustering K in advance. The value of k has a great influence on the positioning accuracy. The method of fingerprint point is still laborious.

There are many ways to get the indoor plane location. In this paper, we use the WiFi fingerprint location algorithm based on weighted k-nearest neighbor (WKNN). WiFi fingerprints from different AP are collected. Due to the correlation of AP and the fluctuation of WiFi signal, the location estimation problem is full of uncertainty. At present, some people have studied the AP selection method. Youssef et al. [8] proposed AP selection strategy based on maximum mean, which selected the AP of maximum mean of WiFi RSS at reference points. This method reduces the effect of weak signal, and obtains better location estimation results with selected APs, but this method does not take into account the regional characteristic and fluctuation of the WiFi signal. Ge et al [9] proposed a distributed strategy of AP selection. The positioning area is divided into some subareas, and then WiFi signal strength values of each AP are in each subarea. Finally, signal strength values to determine the correlation size between each AP and subareas. The method considers the regional characteristic of WiFi signal, but it doesn't consider the fluctuation of WiFi signal. Chen et al. [10] proposed an AP selection method based on information gain, which does not take into account the correlation between AP. Hossain et al. [11] analyzed the effect of the number of AP. The above AP selection methods have a certain improvement on the positioning accuracy, but the methods are not comprehensive enough. In order to get better accuracy, we optimize the method of AP selection.

LBS is mainly focused on two-dimensional positioning in the field of indoor positioning so far. The three-dimensional positioning has not attracted enough attention. Some people have done the research. Yu et al. [12] took the indoor three-dimensional positioning of multi-sensor information fusion as the research goal. He focused on attitude algorithms of inertial sensors, pedestrian dead reckoning (PDR) algorithm and height measurement of barometer. Bal et al. [13] implemented a three-

978-1-5386-3755-5/18/$31.00 ©2018 IEEE

dimensional system of indoor positioning and tracking by ZigBee.

In this paper, according to the difference of WiFi signal strength between floors, the floor is identified by SVM algorithm. The position along the altitude direction is obtained. Then we use the WiFi fingerprint location algorithm based on WKNN to get the indoor plane location. Better location effect is obtained by improved method of the AP selection. Finally the three-dimensional indoor positioning is realized.

The remainder of this paper is organized as follows. In section II, we briefly introduce the principles of SVM. In the section III, we analyze the characteristics of the WiFi signal. In the IV and V sections, we propose AP selection strategy and build the SVM classifier model. In the section VI, we give the experimental results and analysis. Finally, conclusions are presented in section VII.

II. LINEAR SVM PRINCIPLE

SVM is a popular model of binary classification. The model is defined as the linear classifier with the largest interval in the feature space [14-15]. The support vector machine contains the kernel function. This paper utilizes the linear kernel function to support vector machine classification. The specific principle is as follows:

There are a series of training data $(\vec{x}_1, y_1), (\vec{x}_2, y_2), \ldots, (\vec{x}_n, y_n)$, $y_i\,(i = 1, 2, \ldots n)$ equal to 1 or -1 to represent different categories. $\vec{x}_i$ is a vector making up of different features. There are a lot of hyper planes to separate the two types of sample data. In Fig. 1, the pentagram and the circular scatter plot represent two kinds of sample data respectively. Line B and line C are hyper planes, but the maximum gap of the hyper plane is only one, namely optimal hyper plane. SVM uses the maximum distance to find an optimal hyper plane $f(x_i) = \vec{w} \cdot \vec{x}_i + b$. Line C is the best hyper plane in Fig. 1, line A and line D are the extreme hyper plane, line E is the largest hyper plane interval.

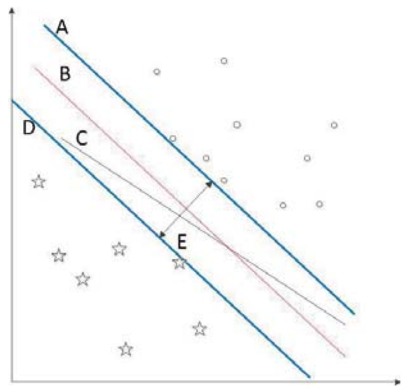

Fig. 1. SVM classification

In the SVM classification method, the maximum geometric interval of the two classes can be weighed by two times of the reciprocal of $\|\vec{w}\|$. The greater interval the smaller $\|\vec{w}\|$. In order to avoid the data falling into the interval region, an additional constraint is added as (1).

$$
\begin{aligned}
y_i = 1 &: \vec{w} \cdot \vec{x}_i + b \geq 1 \\
y_i = -1 &: \vec{w} \cdot \vec{x}_i + b \leq 1
\end{aligned} \tag{1}
$$

Combining the two equations in (1), (2) can be obtained:

$$1 - \mathrm{y}_i \cdot (\vec{w} \cdot \vec{x}_i + b) \leq 0 \tag{2}$$

To avoid the data points falling into the interval region, we ignore a number of data with relatively large deviations. So we consider the loss function as (3):

$$\sum_{i=1}^{n} \max\left(0, 1 - \mathrm{y}_i \cdot (\vec{w} \cdot \vec{x}_i + b)\right) \tag{3}$$

Taking into account the hyper plane interval and loss function, we minimize the (4), and C is a cost parameter[16]:

$$\frac{1}{2}\|\vec{w}\|^2 + C\sum_{i=1}^{n} \max\left(0, 1 - \mathrm{y}_i \cdot (\vec{w} \cdot \vec{x}_i + b)\right) \tag{4}$$

After the minimization (5), the sum of the $\vec{w}$ and b values can be obtained and the (5) can be used for classification.

$$\operatorname{sgn}(\vec{w} \cdot \vec{x}_i + b), \operatorname{sgn}(x) = \begin{cases} -1 : x < 0 \\ 0 : x = 0 \\ 1 : x > 0 \end{cases} \tag{5}$$

III. ANALYSIS OF WIFI SIGNAL CHARACTERISTICS

In order to use the WiFi signal for the right floor recognition and plane positioning, we first need to understand the characteristics of WiFi signal. APs interference by obstacles and multipath effect cause fluctuation of WiFi signal. WiFi signal strength gradually attenuation along with the increasing of distance between AP and collection point, and signal strength will weaken when WiFi signal passes through the wall. These factors may decrease positioning accuracy. We analyze WiFi signal characteristic through the experiment for better deployment of AP and fingerprint next experiments.

A. Analysis of WiFi signal fluctuation

In order to test the fluctuation of WiFi signal, the experiment 1 was carried out in a complex laboratory environment of School Of Geodesy and Geomatics, Wuhan University. The laboratory area is 120m^2. There are 20 desks in the laboratory, and each desk is equipped with computer. There are 2 APs in the room, which are recorded as AP1 and AP2. MI4 smart phone collected the WiFi signal for 8 minutes at each point, and one RSS value is collected every second. In the change curve of AP signal of Fig. 2, we can see that the WiFi signal fluctuate with time, and the changed range of the two AP signal is roughly the same. But the fluctuated frequency of AP1 is relatively high. The fluctuation of AP2 tends to be stable on the whole except big fluctuation at some time. In the probability histograms of Fig. 2, we can see that the distribution of the WiFi signal is similar to the normal distribution, but it does not completely conform to the normal distribution. The distribution of AP1 signal is relatively scattered, while the distribution of AP2 signal is relatively concentrated. The cause of the above results is that the complicated indoor environment leads to multipath. Therefore, when we choose the AP for indoor positioning, we should not only consider the signal strength of the AP, but also consider the stability of the AP signal. In order to get more true values of WiFi RSS in next experiment of floor identification, the RSS data are collected in each point for one minute.

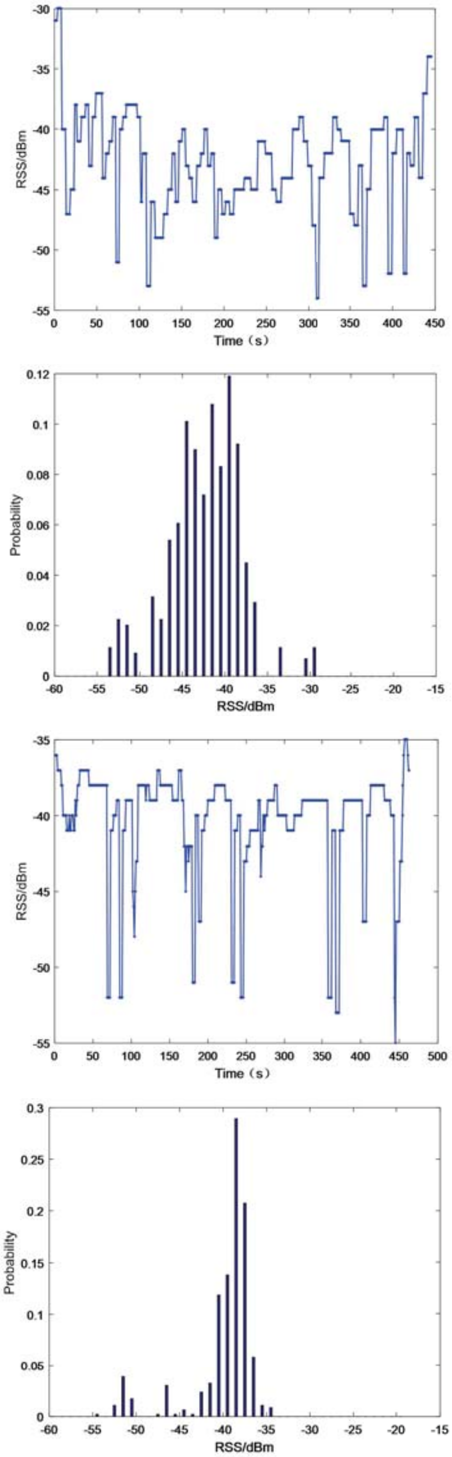

Fig. 2. The curves describe change of AP1 and AP2 signal with the time in constant position. Probability histograms of AP1 and AP2 signal strength show distribution of the RSS. From top to bottom, the first picture and second picture describe the characteristics of AP1. the others describe the characteristics of AP2.

B. The Influence of the Body Block on WiFi Signal

The experiment 2 was carried out in a 50m^2 room. We kept the same position at a collection point. The RSS was received in different directions from same AP, which are the face of AP direction and back directions to AP. The RSS was collected in each direction for 5 minutes, and the histogram of the RSS distribution is shown in Fig. 3. Under the body block, the signal attenuation is more serious. The most concentrated RSS is collection in the face of AP. Therefore, data are collected in all directions to ensure that the training samples are more abundant.

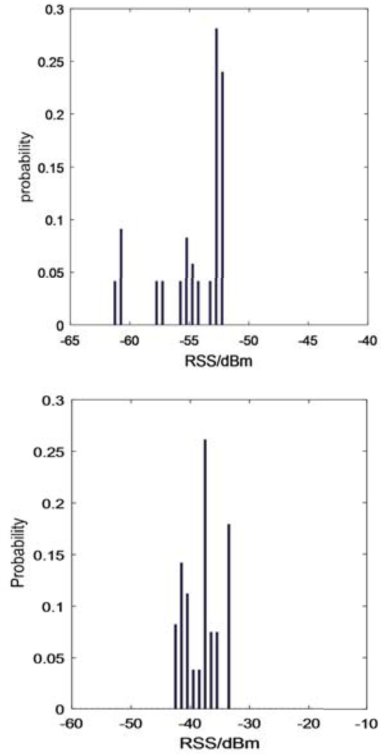

Fig. 3. The histograms describe distribution of RSS values, RSS values are collected in different directions

C. Characteristics of WiFi Signal through Wall

In order to verify RSS change of the AP through wall, we designed the following experiment 3. The experiment describes two cases. One case is that we receive signal from outdoor environment to indoor environment when AP is set outside, and the other is from outdoor environment to indoor environment when AP is in the laboratory In Fig. 4, the change of the RSS is abrupt when people walk into the laboratory or go out of the laboratory. The biggest change of signal strength is about 30dBm, and the average change is about 20dBm. The degree of change can satisfy the identification between floors.

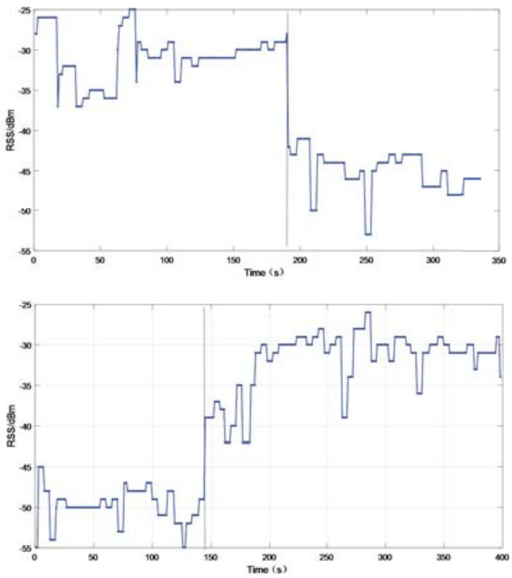

Fig. 4. First picture shows outdoor environment to indoor environment when AP in outdoors environment. Second picture shows outdoor environment to indoor environment when AP in indoors environment

IV. AP SELECTION STRATEGY

According to the characteristics of AP signal, we need consider many factors for AP selection such as signal fluctuation, the regional characteristic of the AP signal, the relevance of AP. In this paper, we propose a synthetical method of AP selection. Firstly, we select uniformly a number of collection points in the positioning area. We collect 60s RSS data and calculate the average RSS value of each AP at each point. We rank average value of each AP in a descending order, and select the Top-N APs from the ranking at each collection point. Then we calculate the standard deviation of each selected AP at each collection point. To filter out the abnormal AP, if the standard deviation is greater than the threshold value of AP, the AP will be removed. Finally, the mutual interference between the APs is considered with the mutual information method. The smaller the mutual information, the smaller correlation between the selected APs. The process of synthetical AP selection is shown in Fig. 5. Through synthetical strategy of above AP selection, the selected APs to estimate location with the WKNN algorithm. This method more fully takes into account the signal characteristics of AP, which gets the optimal selection of AP and improves the accuracy of indoor positioning.

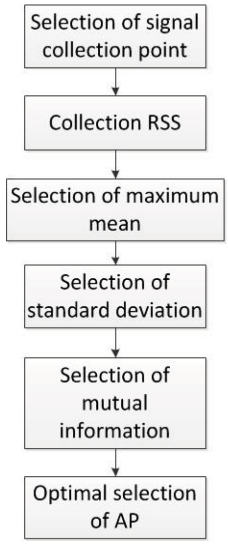

Fig. 5. AP selection process

V. BUILDING PROCESS OF SVM CLASSIFIER MODEL

The classification process is mainly divided into three steps: data acquisition, data samples preprocessing, classifier training and testing. The samples need preprocessing to achieve better classification results. In the training phase of the classifier, we put the samples into two parts, the training data (80%) and test data (20%)[17]. In this paper, we select 8 signal features as one sample to train the classifier. Classifier is evaluated by test data. The detailed process of the establishment of the classifier is shown in Fig. 6.

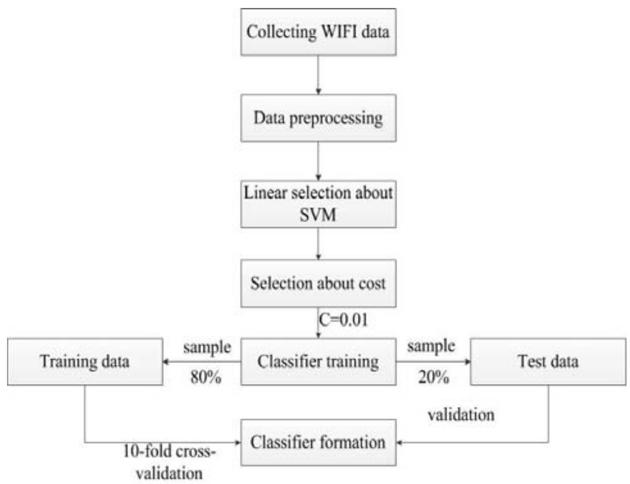

Fig. 6. The process of establishing classifier model

VI. EXPERIMENTAL RESULTS AND ANALYSIS

A. Experimental Deployment

The experiment was carried out in the office building of School Of Geodesy and Geomatics, Wuhan University. The experimental area covered three floors of office buildings. There were three APs on each of the first floor and the third floor. Two APs were deployed on the second floor. The APs of each floor were basically the same position. The position of each AP on the third floor is shown in Fig. 7. The position of red dot represents the position of each AP.

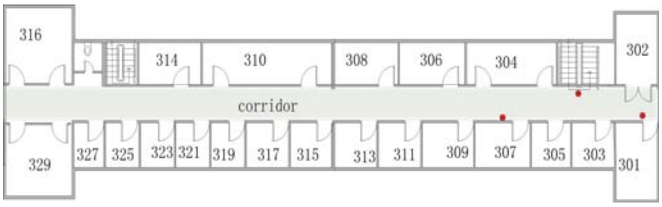

Fig. 7. AP distribution is on the third floor

MI4 smart phone (Android4.4 based MIUI6 operation system) collected data in the corridor, collection position of each floor was basically the same. For example, collection position on the third floor is shown in Fig. 8, the position of the red dot represents the position of each AP, and the blue symbol represents the collection point. The interval between collection points is about 1.2m along the corridor direction, and the interval between collection points is about 0.6m in a direction perpendicular to corridor. 45 collection points are distributed in the corridor. Taking into account the body shelter and signal influence, we collect 30s data at each point and change one direction in each 10 seconds. There is a regular collection along the S route from left to right. The mobile phone is placed horizontally when the data are collected. This method of collection makes the training samples as rich as possible.

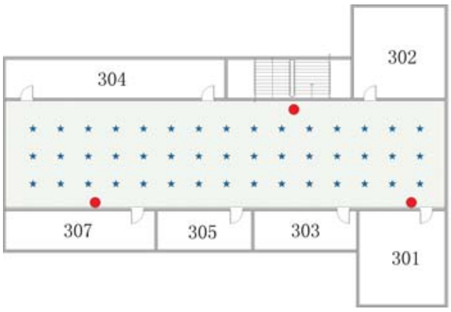

Fig. 8. Fingerprint distribution is on the third floor

In this experiment, we collect RSS among three floors for half an hour on each floor. One sample contains signal features of 8 APs, 1645 samples are collected from the first floor, 1675 samples from the second floor and 1679 samples from the third floor. 4999 samples are collected in total. The number of samples on each floor is as consistent as possible to ensure the reliability of the classification.

In the data collection process, RSS of the AP may not be collected at a certain moment due to the environmental instability. In order to ensure the number of features of each sample, the RSS is replaced by a default value of the -95dBm when the signal can't be received at a time.

In order to get indoor plane location and test the effect of indoor positioning, we did an experiment in a complex laboratory environment of School Of Geodesy and Geomatics, Wuhan University. The laboratory area is 120m^2. There are 20 desks in the laboratory, and each desk is equipped with a computer. MI4 smart phone collected the WiFi signal for one minute at each point, and one RSS value is collected every second. There are students walking back and forth in the experiment. The locations of AP and RP are shown in Fig. 9. The symbol of triangle represents AP, four APs distribute in the four corners of the room. There are multiple APs in the outdoor environment. Mobile phone can receive the WiFi signal from more than ten APs at the same time. The pentagram symbol represents the fingerprint point, which is also frequently called the reference point (RP). The distance between reference points is two meters. The circle symbol presents the location point, which is also called the test point (TP). There are 20 reference points and 10 location points in total in the laboratory. The collection sequence is from the top to the bottom, from the left to the right. These points are numbered in the order of collection.

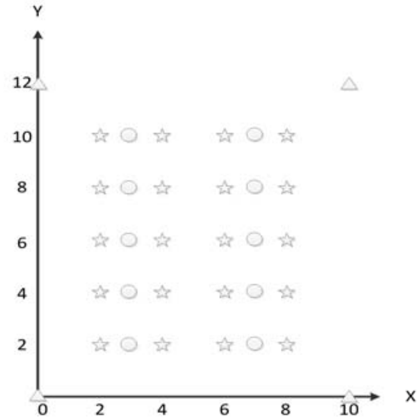

Fig. 9. The distribution of indoor AP, RP and TP point

B. Data analysis and accuracy evaluation

1) Learning process of SVM classifier

Fig. 10 shows the curve of the correct rate of training and cross-validation. In Fig. 10, we can see that the performance of the classifier model on different training sample sets. The two curves converge to about 98.6% with the increase of the training samples. The stable classification results can be obtained at the point of convergence, which can be used to identify the floors well. The curve begins the convergence when the number of samples reaches 1500. The speed of convergence is fast. We take less time to identify the floor than other methods of identification floor.

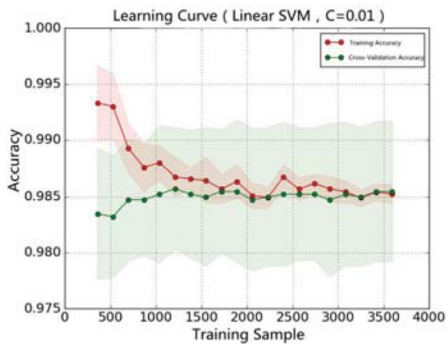

Fig. 10. The curves about the process of learning

2) Evaluation of SVM recognition accuracy

We generally use the accuracy of samples classification to evaluate the effect of SVM classification. By testing the classifier model with the test data (20% of the total samples), we get the confusion matrix of the table 1. The accuracy rate is generally equal to the sum of the diagonal data by the total number of samples in the confusion matrix.

TABLE I. THE CONFUSION MATRIX OF FLOOR IDENTIFICATION OF SAMPLES

Predicted Class / Labeled Class	First Floor	Second Floor	Third Floor
First Floor	330	0	0
Second Floor	0	326	2
Third Floor	4	3	334

In Table 1, we can see that the number of samples of the predicted class and the predicted class is almost the same. The accuracy rate of classification is over 99%, which indicates that the method of floor recognition based on SVM classification can achieve the desired effect. The position along the altitude direction is obtained with WiFi signal.

3) Performance comparison of different AP selection methods

In this paper, the synthetical method of AP selection is proposed. In order to facilitate the record, the method named Union Method. The location effect is tested with the WKNN algorithm. According to the repeated tests with the same other conditions, the positioning effect of K=6 is better in the WKNN algorithm. This paper uses K=6. In Fig. 11, the small circle symbol represents the actual position and the asterisk represents the estimated position. We can see the location effect of synthetical selection method. The average error of each location point is below 1 m.

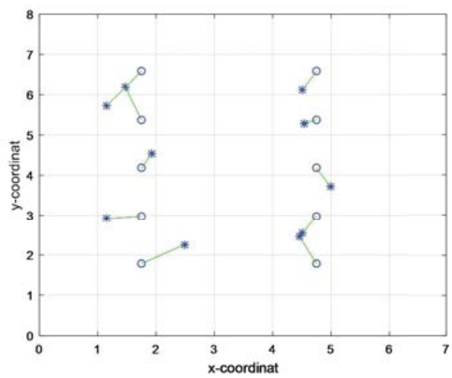

Fig. 11. The comparison of actual position and estimated position

The synthetical method of AP selection is compared with other methods, such as the maximum mean method (MM), information gain (IG) and so on. We find that synthetical method of AP selection can get better positioning effect in Fig. 12. The average error of the synthetical method is above 0.4 m smaller than other methods. The method of maximum mean doesn't take into account the fluctuation of WiFi signal and correlation of APs. The information gain method doesn't consider the correlation of APs. But the proposed method of this paper considers above the characteristics of WiFi signal.

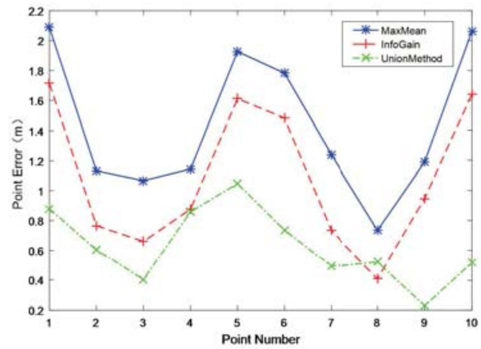

Fig. 12. Performance comparison of different AP selection methods

4) Three-dimensional positioning effect

Through the floor recognition, we obtain the position along the altitude direction. And we get the indoor two-dimensional positioning with the WiFi fingerprint algorithm based on WKNN. At last, we realize the three-dimensional positioning of indoor environment. In Fig. 13, we show the effect of three-dimensional positioning.

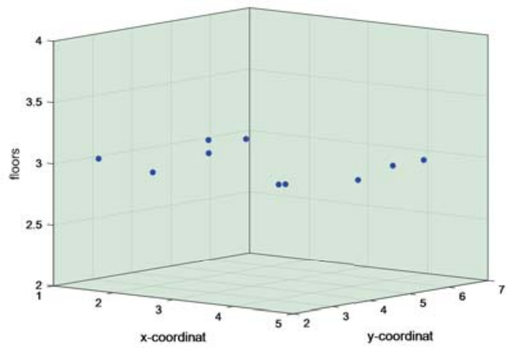

Fig. 13. The effect of three-dimensional positioning

VII. CONCLUSIONS

In this paper, we propose a method of floor identification based on SVM. The method uses the WiFi signal of each floor to identify the floor by SVM, and which obtains the position along the altitude direction. The method utilizes the linear kernel function for SVM classification and the 10-fold cross-validation to train classifier model. The positioning method costs less time and gets high accuracy. The accuracy rate of floor recognition is 99.09%. In addition, we improve the AP selection method based on the existing method. The improved method includes maximum mean value of WiFi RSS calculation, standard deviation filtering and mutual information analysis. It considers the signal fluctuation, regional characteristic and AP correlation. The improved method achieves optimization of AP selection. Through the improved method of AP selection and the WiFi fingerprint algorithm based on WKNN, we acquire the better accuracy of plane positioning. The average error of location point is above 0.4 m smaller than other methods. Finally, the indoor 2.5D positioning is realized by the floor recognition and the plane positioning. However, the paper only realizes three-dimensional positioning through the algorithm, the next step will be to achieve the real-time 2.5D positioning system.

ACKNOWLEDGMENT

This research is sponsored by National Natural Science Foundation of China (No.41474004). The work is also supported by funds from the Key Laboratory of Precise Engineering and Industry Surveying, NASMG. The funding project number is PF2017-7. Thanks to research Center for High Accuracy Location Awareness to provide the help. Heartfelt thanks to the editors and anonymous referee experts provide valuable opinions and suggestions for this article.

REFERENCES

[1] R. Xi, Y. Li, M. Hou, "Indoor positioning method on wifi," Computer Science, vol. 42, no. 2, pp.1-4, 2017.

[2] G. Shi, B. Wang, B. Wu, "An overview of the indoor location methods based on WiFi and mobile intelligent terminals," Computer Engineering, vol. 41, no. 9, pp.39-44, 2015.

[3] P. A. Karegar, "Wireless fingerprinting indoor positioning using affinity propagation clustering methods," Wireless Networks, pp.1-9, 2017.

[4] I. Bisio, M. Cerruti, F. Lavagetto, M. Marchese, M. Pastorino & A. Randazzo, "A trainingless wifi fingerprint positioning approach over mobile devices," IEEE Antennas & Wireless Propagation Letters, vol. 13, no. 1, pp. 832-835, 2014.

[5] S. Li, H. Li, & L. Sun, "Privacy-preserving crowd sourced site survey in wifi fingerprint-based localization," Eurasip Journal on Wireless Communications & Networking, 2016.

[6] H. Li, D. Qi, Q. Wang, "A floor discrimination method based on the RSSI difference between floors", CN103916821A, 2014.

[7] Z. Deng, W. Wang, L. Xu, "A WLAN indoor positioning method of floor identification based on K-means algorithm," Software, vol. 33, no. 12, pp.114-117,2013.

[8] M. A. Youssef, A. Agrawala, A. U. Shankar, "WLAN location determination via clustering and probability distributions," IEEE International Conference on Pervasive Computing and Communications, pp.143-150, 2013.

[9] L. Ge, X. Zhao, K. Li, "Application of. AP distributed sensor selection strategy in the indoor location," Transducer and Microsystem Technologies, vol. 34, no. 9, pp.154-157, 2015.

[10] Y. Chen, Q. Yang, J. Yin "Power-efficient access-point selection for indoor location estimation," IEEE Transactions on Knowledge & Data Engineering,vol. 18, no. 7, pp.877-888, 2006.

[11] A. K. Hossain, W. S. Soh, "Bound Analysis of Localization Using Signal Strength Difference as Location Fingerprint," Conference on Information Communications, pp.1981-1989, 2010.

[12] Yu Qianyong. "Multi-sensor information fusion Indoor 3D location algorithm and implementation," China University of Mining and Technology, 2015.

[13] M. Bal, H. Xue, W. Shen, "A 3-D indoor location tracking and visualization system based on wireless sensor networks," IEEE International Conference on Systems Man and Cybernetics, pp.1584-1590, 2010.

[14] H. Qian, Y. Mao, W. Xiang, & Z. Wang, "Recognition of human activities using svm multi-class classifier," Pattern Recognition Letters, vol. 31, no. 2, pp.100-111, 2010.

[15] A. Palaniappan, R. Bhargavi, & V. Vaidehi, "Abnormal human activity recognition using SVM based approach," International Conference on Recent Trends in Information Technology, pp.97-102, 2012.

[16] V. Cherkassky, Y. Ma "Practical selection of svm parameters and noise estimation for svm regression," Neural Networks the Official Journal of the International Neural Network Society, vol. 17, no. 1, pp.113-126, 2004.

[17] R. Yasmine, P. Ling, "Indoor fingerprinting algorithm for room level accuracy with dynamic database," Fourth International Conference on Ubiquitous Positioning, Indoor Navigation and Location Based Services, pp.113-121, 2017.