Fingerprint Localization Algorithm Based on Occlusion Judgment

Haoyang Qi, Zhigang Liu, Dong Liu, Lei Wang

School of Computer and Communication Engineering, Northeastern University, Qinhuangdao
E-mail: zliu@neuq.edu.cn

Abstract: With the development of Device-free localization technology, indoor fingerprint localization has attracted more and more attention. However, in indoor environment, the degree of occlusion at different locations is obviously different, the positioning accuracy based on Channel State Information (CSI) system decreases. In this paper, we propose a fingerprint localization algorithm based on occlusion judgment (FLOJ algorithm). In the fingerprint processing phase, we use principal components analysis (PCA) to reduce the dimension of fingerprint database to remove redundant data. And k-nearest neighbor (KNN) position estimation algorithm is used to get the matching results. In the phase of multi-position fusion, we use the algorithm based on CSI propagation model as the criterion to measure the occlusion degree of indoor environment. And we fuse the matching positions obtained by multi-receivers at the same time. Experiment results show that compared to KNN localization methods, our proposed algorithm can effectively reduce positioning error.

Key Words: Wi-Fi localization, Channel State Information, Occlusion Judgment, Multi-location fusion.

1. INTRODUCTION

With the rapid development of intelligent home system, more and more researchers have invested in indoor localization. Many techniques for indoor localization have been proposed so far. With the popularity of Wi-Fi in our daily life, the effects of human behavior on surrounding wireless signals can be exploited as a powerful inductive tool for locating and identifying the activities of people in the vicinity of the wireless network. It can help the wireless transceiver to obtain accurate measurements. This emerging technology is called as Device-free localization (DFL). Because it does not require anyone to be tested to wear any equipment, it can be better applied to many key technologies in potential applications.

Indoor positioning can be classified according to the equipment used [1], [2]. The most common methods include camera, infrared, ultrasonic, RFID and Bluetooth. However, the traditional methods have many disadvantages. The method based on the wearable sensor is inconvenient, and the elderly always forget to wear their sensors. Camera-based localization has strict requirements for lighting and line-of-sight conditions and may result in privacy leaks. Radar-based methods are expensive and have limited coverage, and cannot meet high-precision positioning in indoor environment. The client can receive and process the data packet to obtain the wireless signal. According to the positioning methods, we can roughly divide the localization into three categories:1) positioning can be achieved by simulating the propagation of the wireless signal [3]. 2) We also can use the recognition algorithm to establish fingerprint database to achieve positioning [4], [7]. 3) Some researchers even can use the knowledge of image processing to transform the information into an image and process the image to achieve positioning [6], [8]. However, RSS obtained with current WIFI devices is less accurate, but RSS information can also be used for some applications, such as indoor target motion detection [9]

This paper uses WIFI devices to locate indoor target. One of the main advantages of using Wi-Fi is its low cost, which can be used by different types WIFI without additional hardware. Another advantage of WLAN is that it is not necessary to maintain line of sight conditions (LOS). The most widely used WLAN localization method is to utilize RSSI and CSI signals that are easily extracted in an IEEE 802.11 network site.

If there are multiple acquisition cards to collect WIFI information in the room, in other words, there will be multiple location results. So how to fuse effectively is a problem worth solving. We propose an indoor localization system based multiple receivers. Specifically, the CSI propagation model is used as the standard of indoor environmental interference, and the localization results from multiple paths are effectively fused to improve the positioning accuracy.

The rest of this paper is structured as follows. The section 2 introduces the relevant knowledge of CSI. In section 3, the system architecture and basis concepts. In Section 4 FLOJ algorithm is introduced. Section 5 validates the proposed scheme through experimental evaluation. Finally, the conclusion is drawn in section 6.

2. PROCESSING OF CSI DATA

2.1 Basic Concept of CSI

CSI is the abstract of Channel State Information, which belongs to the PHY layer information. By using an OFDM to modulate signal, CSI provides a fine-grained physical information, more sensitive to the environment, so it is used

This research is supported by the National Nature Science Foundation of China under Grant 61473066, and the Fundamental Research Funds for the Central Universities under Grant N152302001.

978-1-7281-0106-4/19/$31.00 ©2019 IEEE

in motion recognition, indoor accurate positioning and tracking and other fields.

After emitted by the transmitter, the signal is attenuated by the refraction and reflection of the indoor environment, the received signal at the receiver can be expressed as:

$$\mathrm{Y} = \mathrm{H} \cdot \mathrm{X} + \mathrm{N} \tag{1}$$

where X and Y represents the transmitter and receiver signals. N represents the noise, and H represents the CSI complex matrix. After OFDM modulation, a single subcarrier can be represented by amplitude and phase.

$$H_k = ||H_k|| \cdot e^{j\angle\theta} \tag{2}$$

where $||H_k||$ represents the amplitude of H_k and θ represents the phase of H_k. Specially, this paper uses CSI amplitudes $||H_k||$ to generate fingerprint databases.

2.2 CSI time domain transform

We collect data in the frequency domain through the acquisition card. In order to make full use of the CSI information, we can convert it to Channel Impulse Response (CIR) signal by Fourier transform. CIR signal can be expressed as:

$$\mathrm{h}(\tau) = \sum_{i=1}^{N} \alpha_i e^{-j\theta} \delta(\tau - \tau_i) \tag{3}$$

where α_i,θ_i and τ_i are the amplitude, phase, and time delay of the i_{th} path, respectively. N is the total number of multipath and δ(τ) is the Dirac delta function. Each impulse represents the certain multipath component multiplied by the corresponding amplitude and phase, respectively.

2.3 CSI Propagation Model

For CIR data, a weighting method for 30 subcarriers in frequency band on the central frequency is proposed in Reference [1]. The specific formula is given as:

$$CSI_{eff} = \frac{1}{K}\sum_K \frac{f_k}{f_c} \times ||A||_k \tag{4}$$

where CSI_{eff} is the input for propagation model, K is the sum of numbers of subcarriers. $||\mathrm{A}||_k$ is the amplitude of k_{th} subcarrier. And then the distance can be calculated as:

$$\mathrm{d} = \frac{1}{4\pi}\left[\left(\frac{v}{f_c \times CSI_{eff}}\right)^2 \times \sigma\right]^{\frac{1}{n}} \tag{5}$$

where v is the velocity of the transmitted wave, n is the attenuation factor, and σ denotes all other hardware factors including transmitted power, antenna gains, and so forth. N and σ need to be fitted for specific data processing because they are related to specific indoor environment.

3. SYSTEM ARCHITECTURE

3.1 Structure Description

In this section, we first give an overview of our localization system, and then describe the key steps of the proposed localization algorithm.

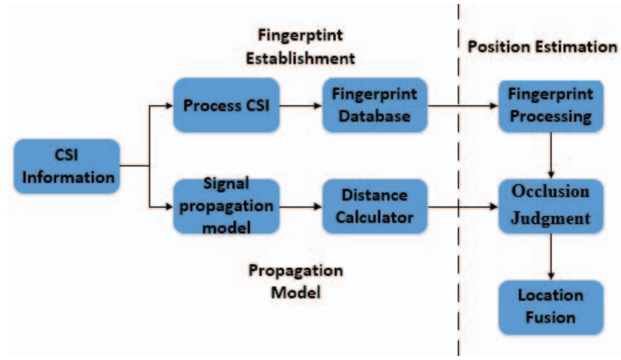

Fig.1 System architecture

Fingerprint establishment: In the fingerprint database establishment phase, we measure the CSI frequency data of each reference point. If there are multiple receivers, store their data in sequence in the fingerprint database.

Propagation Model: Performing a Fourier transform on the CSI data of each point to convert the frequency domain signal of the CSI into a time domain signal. Then each packet is weighted by formula (1), finally the unknown environmental parameters n and σ are fitted with the input data.

Position estimation: In this phase, it is divided into three parts. Firstly, the fingerprint database is used to calculate a fingerprint position using KNN algorithm. Then the propagation model method is used to calculate the location of a model propagation method. Finally, the propagation model is used as a standard to judge the occlusion degree, and the fingerprint location method is fused by weighting.

3.2 CSI Fingerprint Algorithm

Next, the fingerprint localization process will be described. Firstly, the fingerprint database is established. In our test environment, there are three Intel 5300 wireless network cards, each card can collect 30 subcarriers. It can be described as:

$$F = [f_1^1, f_2^1, \ldots, f_{30}^1, f_1^2, f_2^2, \ldots, f_{30}^2, f_1^3, f_2^3, \ldots, f_{30}^3] \tag{6}$$

In this way, we have established the fingerprint database of the indoor environment in this paper. Similarly, the test points are sampled in the same way, and the test data can also be obtained as follows:

$$T = [t_1^1, t_2^1, \ldots, t_{30}^1, t_1^2, t_2^2, \ldots, t_{30}^2, t_1^3, t_2^3, \ldots, t_{30}^3] \tag{7}$$

In this paper, KNN is used as a pattern recognition method to realize localization. The theory is that in feature space, if most of the k similar samples of a sample belong to a certain category, then the sample also belongs to this category. The measure of similarity is the Euclidean distance calculated between train and test samples. The smaller the Euclidean distance, the higher the similarity. The Euclidean distance is defined as:

$$\mathrm{l(f,t)} = \sqrt{\sum_{i=1}^{3}\sum_{j=1}^{N}(f_j^i - t_j^i)^2} \tag{8}$$

We set the location by using fingerprint algorithm as (x_k, y_k).

4. FLOJ ALGORITHM

4.1 Professing CSI Data

In the fingerprint algorithm we need to deal with 30*3 - dimensional data samples. Such data has the obvious drawback that redundant parts of the data may affect the localization results. We have to separate the valid part of the data. Data dimension reduction is an effective method. In this paper, we use Principal Component Analysis (PCA) to process the data. The principles of PCA are briefly described below:

Firstly, we let X donate the train data of CSI, where F_i is one data with 30 subcarriers of the train database. M is the number of train data.

$$X = [F_1'; F_2'; \ldots; F_k'; \ldots; F_m'] \tag{9}$$

Secondly, we centralize X. Calculate the average of each column, donated as:

$$\bar{X} = [\overline{F_1}'; \overline{F_2}'; \ldots; \overline{F_m}'] \tag{10}$$

$$X_{new} = X - \bar{X} = [F_1' - \overline{F_1}'; \ldots; F_k' - \overline{F_k}'; \ldots; F_m' - \overline{F_m}'] \tag{11}$$

The covariance matrix of X_{new} is calculated and we denoted as cov_w .Meanwhile, the cov_w eigenvalues and eigenmatrices are also calculated as U.

And then the eigenvalues are sorted in descend order, the largest k of eigenvalues are selected, and then the corresponding k eigenvectors U are respectively used as column vectors to form the eigenvector matrix W_{new}.

Finally, we Calculate $W_{new}^T \cdot X_{new}$, which is to project the dataset X_new onto the selected feature vector, and thus we get reduced-dimensional dataset $W_{new}^T \cdot X_{new}$.

4.2 Occlusion Judgment

We can get the data collected by three sets of data acquisition cards through fingerprinting, but we can not simply judge whether the three sets of data are good or bad. Averaged localization results will introduce the results with large errors. So it is very important to find a reasonable weighted fusion algorithm. Wireless signal propagation model can accurately calculate the distance from the location result (x_k, y_k) to the WIFI when the LOS path has less occlusion. If the indoor occlusion is too large in the propagation path, the localization results will be distorted. Thus this paper uses the wireless signal propagation model as the standard to measure the indoor occlusion.

In the condition of single receiver, we set the position of transmitter as (x_p, y_p) the specific principle is introduced as follows, we take WIFI as the center of the circle and d_m as the radius calculated by the propagation model. Then the fingerprint location result (x_m, y_m) and WIFI are connected, the intersection of the them is assumed to be the location calculated by the propagation model method. We set that position as (x_k, y_k). We can get the locations of propagation model method separately if there are multiple receivers in the environment. Then we can compute the Euclidean distance between the (x_k, y_k) and the (x_p, y_p).

$$d = \sqrt[2]{(x_p - x_k)^2 + (y_p - y_k)^2} \tag{12}$$

If the result is smaller, the localization result is more reliable. According to this criterion, we can occlusion judgment on the positioning results of multiple receivers.

4.3 Location Fusion

We assumed that there are j acquisition cards in the indoor environment, and according to the theory described earlier, we write the (x_k, y_k) and (x_p, y_p) in two matrices, described as:

$$P = \begin{bmatrix} x_p^1 & y_p^1 \\ \vdots & \vdots \\ x_p^i & y_p^i \\ \vdots & \vdots \\ x_p^j & y_p^j \end{bmatrix} \qquad K = \begin{bmatrix} x_k^1 & y_k^1 \\ \vdots & \vdots \\ x_k^i & y_k^i \\ \vdots & \vdots \\ x_k^j & y_k^j \end{bmatrix} \tag{13}$$

According to the formula (12) above, we calculate the Euclidean matrix D. and then use the reciprocal of D as the weight W. And then we normalize W. However, we know that although the introduction of weights will improve the accuracy, the positioning results with large errors will affect the positioning accuracy more or less. So we decided to set a threshold and filter out the result that the weight is less than the threshold, which will further improve the positioning accuracy. Finally, the multi-receive positioning result can be expressed as:

$$D = [d_1, d_2, \ldots, d_k \ldots d_j] \tag{14}$$

$$W = \frac{1}{D} \tag{15}$$

$$X = \frac{W' * x_m}{\sum_{i=1}^{l} ||W||} \qquad Y = \frac{W' * y_m}{\sum_{i=1}^{l} ||W||} \tag{16}$$

5. EXPERIMENTS VALIDATION

5.1 Environment Layout

We consider the rectangular area shown in Figure 2, which represents the indoor environment, as the localization area. The area is divided into M×N grids, with a datum point set in the upper left corner of the grid. In the positioning area, there is one WIFI as the transmitter and three acquisition cards as the receiver. WIFI sends signals to the receiver all the time. Their positions are known, but there are several obstacles between them that are unknown and interfere with signal propagation. There is only one user in an unknown location throughout the region. We validate our theory in a typical indoor scene. As shown in Figure 2, the entire experimental area is approximately 6m×4.8 m. In this paper, we only focus on positioning in 2D space. We select 20 reference points and 9 test points. The size of each square is $0.6 \times 0.6m^2$.

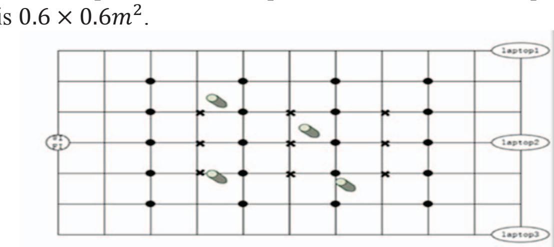

Fig .2 Layout of the room

We use a Dell laptop equipped with Intel 5300 wireless network card as the target device. TP-link TL-WR886N wireless router which has 3 antennas is used as the AP. Another PC trains these data. During the training phase, we collected 2 minutes at each reference point and collected 200 packets. During the testing phase, we collected 30 seconds at each reference point and collected 25 packets. Figure 3 is the CSI data collected at a certain moment.

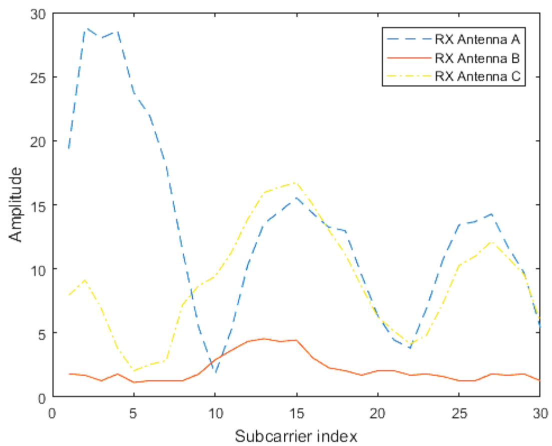

Fig.3 Three Antenna CSI Amplitude

5.2 Comparison with KNN Algorithm

The goal of this paper is to achieve the fusion of multi-receiver location results, so we compare the improved algorithm with the multi-position averaged algorithm. For convenience, we call this algorithm FOJ algorithm. First, the RMSE of the improved algorithm and multi-position averaged algorithm is shown in the following table,

Table 1 Positioning Error (RMSE)

Algorithm	KNN	FLOJ
RMSE(m)	1.019	0.894

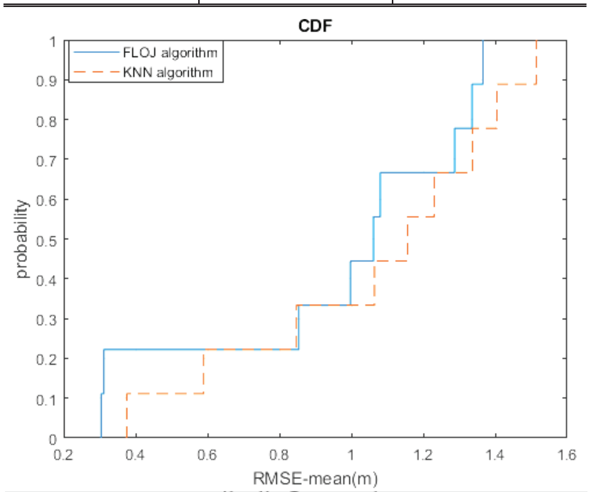

Fig.4 CDF of RMSE with two algorithms

In order to reflect the repeatability of the experiment, we calculated the average errors of 9 locations of the test data, each of which contains 25 data packets. Then the cumulative distribution function (CDF) curves of the two methods are plotted. From the image, we can see that the performance of the improved algorithm has been greatly improved.

6. CONCLUSIONS

This paper presents an indoor Wi-Fi positioning system based on multiple receivers. We construct fingerprints with CSI values and further process them with PCA algorithm to reduce their dimensions. Then we use KNN to obtain the estimated target location. In particular, we use the wireless propagation model as a measure of interference to fuse the location results of multiple receivers.

Through experiments, we validate the proposed indoor localization algorithm. The experimental results show that our algorithm is better than KNN algorithm mentioned in [6]. Compared with KNN algorithm, the positioning accuracy can be improved by 10%, and can obtain more stable target location.

REFERENCE

[1] Mainetti. Luca, L. Patrono, and I. Sergi, A survey on indoor positioning systems, International Conference on Software IEEE, 2015.

[2] S. N. He, S.-H. Gary Chan, Wi-Fi Fingerprint-based Indoor Positioning: Recent Advances and Comparisons, IEEE Communications Surveys & Tutorials (2015):1-1.

[3] Ossi Kaltiokallio, Huseyin Yigitler and Riku Jantti, A Three-State Received Signal Strength Model for Device-Free Localization. IEEE Transactions on Vehicular Technology, 2017 OCTOBER, VOL. 66, NO. 10,

[4] Mager Brad, Philip Lundrigan, and Neal Patwari, Fingerprint-Based Device-Free Localization Performance in Changing Environments, IEEE Journal (2015):2429-2438.

[5] H. Chen, Y.F. Zhang, W. LI, X. F. Tao, and Ping. Zhang, ConFi: Convolutional Neural Networks Based Indoor Wi-Fi Localization Using Channel State Information, IEEE Access (2017) PP.99.

[6] Q. W. Song, S. T Guo, X Liu, and Y. Y Yang, CSI Amplitude Fingerprinting Based NB-IoT Indoor Localization, IEEE Internet of Things Journal (2018):1494-1504.

[7] Q.H. Gao, et al. CSI-based Device-free Wireless Localization and Activity Recognition Using Radio Image Features, IEEE Transactions on Vehicular Technology (2017):1-1.

[8] X. Y Wang, L. J. Gao, S. W. Mao and Santosh Pandey, CSI-Based Fingerprinting for Indoor Localization: A Deep Learning Approach IEEE Transactions on Vehicular Technology (2017):763-776.

[9] R. Ma, Q. Guo, C. Z. Hu and J. F. Xue, An Improved WiFi Indoor Positioning Algorithm by Weighted Fusion, sensors (2015):21824-21843.

[10] Y Zheng, Z. Z. Zhou, Y. H. Liu, From RSSI to CSI: Indoor Localization via Channel Response, ACM Computing Surveys, 2013, Vol. 46. No. 2, Article 25

[11] W. X. Xue, X. H. Hua, Q. Q. Li, Kegen Yu, and W. N Qiu, C Improved Neighboring Reference Points Selection Method for WIFI-Based Indoor Localization, IEEE Sensors Letters (2018):1-1.

[12] L. S. Zhao, H. P. Wang. P. P. Li, J. T. Liu, An improved WiFi indoor localization method combining channel state information and received signal strength, 2017 Chinese Control Conference IEEE.

[13] Q. Lei, H. J. Zhang, H. Sun, and L. L. Tang, Fingerprint-Based Device-Free Localization in Changing Environments Using Enhanced Channel Selection and Logistic Regression, IEEE Access 6(2018):2569-25