

RESEARCH

Open Access



Research on Wi-Fi indoor positioning in a smart exhibition hall based on received signal strength indication

Qing Yang^{*} , Shijue Zheng, Ming Liu and Yawen Zhang

Abstract

To improve the management of science and technology museums, this paper conducts an in-depth study on Wi-Fi (wireless fidelity) indoor positioning based on mobile terminals and applies this technology to the indoor positioning of a science and technology museum. The location fingerprint algorithm is used to study the offline acquisition and online positioning stages. The positioning flow of the location fingerprint algorithm is discussed, and the improvement of the location fingerprint algorithm is emphasized. The raw data of the RSSI (received signal strength indication) is preprocessed, which makes the location fingerprint data more effective and reliable, thus improving the positioning accuracy. Three different improvement strategies are proposed for the nearest neighbor classification algorithm: a balanced joint metric based on distance weighting and a compromise between the two. Then, in the experimental simulation, the positioning results and errors of the traditional KNN (k-nearest neighbor) algorithm and three improvement strategy algorithms are analyzed separately, and the effectiveness of the three improved strategy algorithms is verified by experiments.

Keywords: Indoor positioning, Wi-Fi, Location fingerprint, Nearest neighbor classification, Received signal strength indication

1 Introduction

Science and technology museums are public science education institutions with exhibition education as its main function. There, scientific interest is stimulated, and scientific concepts are enlightened by participation, experience, interactive exhibits and auxiliary displays. For the science and technology museum, it is necessary to know the real-time specific orientation of each visitor in the exhibition hall and the number of visitors and their residence time for each booth. Moreover, according to the number of visitors to each booth, the management of the exhibition center can maintain the better order.

At present, there are a variety of indoor positioning technologies [1], which are broadly divided into the technologies of Wi-Fi positioning [2], radio frequency identification (RFID) positioning [3], ZigBee positioning [4], Bluetooth positioning [5], ultra-wide-band (UWB)

positioning [6], infrared positioning [7], ultrasonic positioning [8], optical tracking positioning [9], and computer vision positioning [9]. Among them, the more influential positioning systems are: the active badge positioning system based on infrared positioning technology researched by the Cambridge Laboratory [10]; the RADAR positioning system [11] realized by Wi-Fi positioning technology in wireless LANs and the Active Bat positioning system [12] realized by ultrasonic positioning technically researched by Microsoft Design Institute; the Cricket positioning system [13] researched by MIT which is an improvement on the Active Bat system; AHLoS (ad hoc localization system) positioning system [14] researched by UCLA which is an improvement on the Cricket system.

There are many different exhibition halls in science and technology museums. Each exhibition hall has many exhibitions. The managers of science and technology museums want to know the flow of people in different exhibitions at different times. Human flow data are conducive to scientific management. With the development

^{*} Correspondence: yangqing@mail.ccnu.edu.cn
School of Computer, Central China Normal University, Wuhan 430079,
People's Republic of China

of the mobile Internet, Wi-Fi hotspots have been covered in most public indoor areas of the city, and Wi-Fi wireless LANs have data communication functions. Wi-Fi wireless networks are easy to deploy and have good scalability. However, accuracy still needs to be further improved. Indoor positioning technology based on the Wi-Fi mobile terminal has a great advantage in positioning in the exhibition hall of science and technology museums. The purpose of this paper is to improve the indoor location algorithm in the environment of science and technology museum, to improve the accuracy of indoor location, and to minimize the complexity of the algorithm, so that it can be used in small- and medium-sized science and technology museums or museums. Based on the literature surveys about the various existing Wi-Fi indoor positioning algorithms and analyses of their advantages and disadvantages, an effective location fingerprint algorithm is proposed. Performance analysis of the proposed algorithm and an evaluation of the algorithm with respect to other existing algorithms are given. This improved algorithm can accurately describe the flow data of people in science and technology museums when it is used for indoor positioning in science and technology museums.

The rest of this paper is organized as follows. Section 2 discusses related work, followed by the improvement of the location fingerprint algorithm designed in Section 3. Section 4 shows the simulation experimental results, and Section 5 concludes the paper with a summary and future research directions.

2 Related work

This section provides a brief review of the common Wi-Fi indoor positioning algorithms. After comparing these algorithms, the location fingerprint algorithm is selected for the indoor location of science and technology museum. The architecture of the Wi-Fi location is introduced, and the flow of the location fingerprint algorithm is described.

2.1 Wi-Fi indoor positioning algorithm

Wi-Fi indoor positioning algorithms [15–17] are commonly found as follows: (1) based on the central point algorithm, (2) the propagation model algorithm, and (3) the location fingerprint algorithm.

The location fingerprint algorithm is simpler and easier to implement on the mobile terminal, and the location accuracy based on location fingerprint feature matching is very robust. Wi-Fi fingerprint positioning has become the preferred technology for indoor environment positioning and navigation because of its convenient use of existing Wi-Fi network infrastructure, low cost and strong environmental adaptability. Therefore, the Wi-Fi location fingerprint algorithm based on a mobile terminal is selected

to conduct experiments in a science and technology museum.

The important task of signal-based fingerprint location is to select the appropriate location algorithm. There are three typical methods: (1) neighbor method; (2) probability method; (3) neural network method. The number of hidden layers of the neural network method can be one or more. The more hidden layers, the higher the accuracy of data fitting, but the computational complexity and training time will also be improved. Normal servers cannot afford the computational load of the neural network method [18]. This method is not suitable for small- and medium-sized science and technology museums. The main work of this paper is to study the indoor positioning method suitable for small- and medium-sized science and technology museums.

2.2 Wi-Fi positioning architecture

In a Wi-Fi indoor positioning system, the position of the reference point is known, and the position of the target node is usually the one to be tested. The mobile terminal is at the location to be tested. During the positioning process, all the surrounding Wi-Fi signal source APs are perceived, and the metrics are uploaded (the specific metrics may be the time of arrival (TOA) of the signal, or the angle of arrival (AOA) of the signal, the time difference of arrival (TDOA) of the signal and received signal strength RSSP.) to the positioning server background. The server background estimates the position coordinates of the point to be tested by a specific positioning algorithm, and then sends it to the terminal and displays it, as shown in Fig. 1.

In a Wi-Fi wireless network environment, a Wi-Fi signal access point (i.e., an AP) periodically broadcasts a beacon signal frame, and although a mobile terminal device with a Wi-Fi access function module does not establish a connection with any AP, it can also receive at least three important parameter indicators of the transmitting source AP from the AP broadcast signal frame: the Wi-Fi AP's MAC address (BSSID), the Wi-Fi AP name (SSID), and the Wi-Fi signal strength indication (RSSI). These three indicators provide strong support for the model for positioning based on location fingerprint feature matching.

In general, the following factors affect the Wi-Fi positioning effect [19]: the absorption effects of Wi-Fi signals, the reflection and diffraction of Wi-Fi signals, the multipath propagation and shadow attenuation of Wi-Fi signals, the size of indoor areas, the changes in temperature and humidity in indoor environment and the signal interference from other electronic devices. In addition, due to the lack of uniform rules and optimization of Wi-Fi, the network environment of the Wi-Fi AP is extremely complicated, and the number of manufacturers producing Wi-Fi

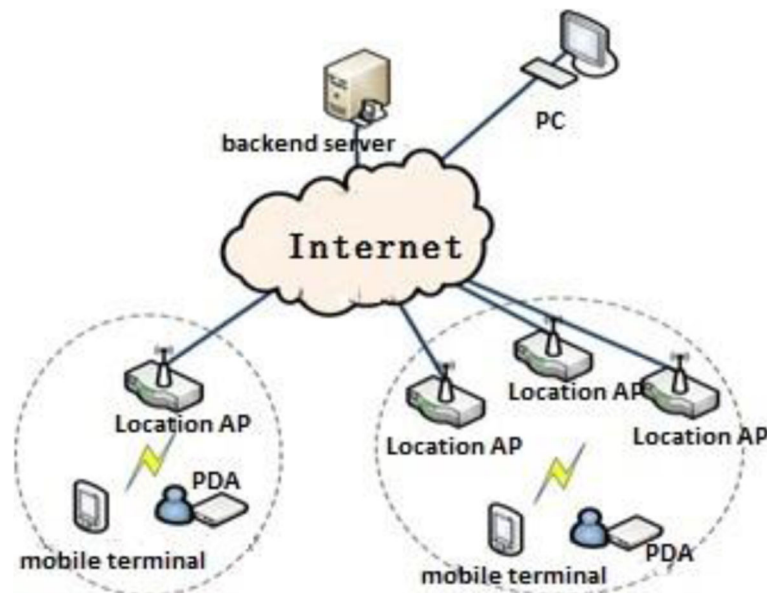


Fig. 1 Overall architecture of Wi-Fi positioning

APs and terminals is large, so the performance difference between devices is very obvious, which means that the various influencing factors must be considered in a specific position to improve the accuracy of positioning.

2.3 The location progress of the position fingerprint algorithm

The position fingerprint algorithm is a feature matching algorithm, and the feature value is the position fingerprint data of the Wi-Fi signal strength.

In the algorithm based on the propagation model, since the Wi-Fi signal is affected by various factors in the indoor environment, the attenuation often does not meet the theoretical value; a position fingerprinting algorithm based on feature matching is proposed. That is, the signal strength information in a group of different Wi-Fi APs is received at each position, and the fingerprint data vector signature sequences of all APs generated by different positions make differences. In this way, specific Wi-Fi position fingerprint information can be obtained at each location. The overall process of position fingerprint algorithm location is shown in Fig. 2.

As Fig. 3 shows, the locating process has two stages: the offline acquisition stage and the online positioning stage. The main task of the offline acquisition stage is to collect the fingerprint data of the reference point position in the indoor environment where the reference point has been deployed, and build a fingerprint database. In the online positioning stage, the fingerprint data collected by the node to be tested is matched with the fingerprint database, and a matching algorithm is used

to find a group of data that best matches it, thereby finding the position of the node to be tested.

2.3.1 Offline acquisition stage

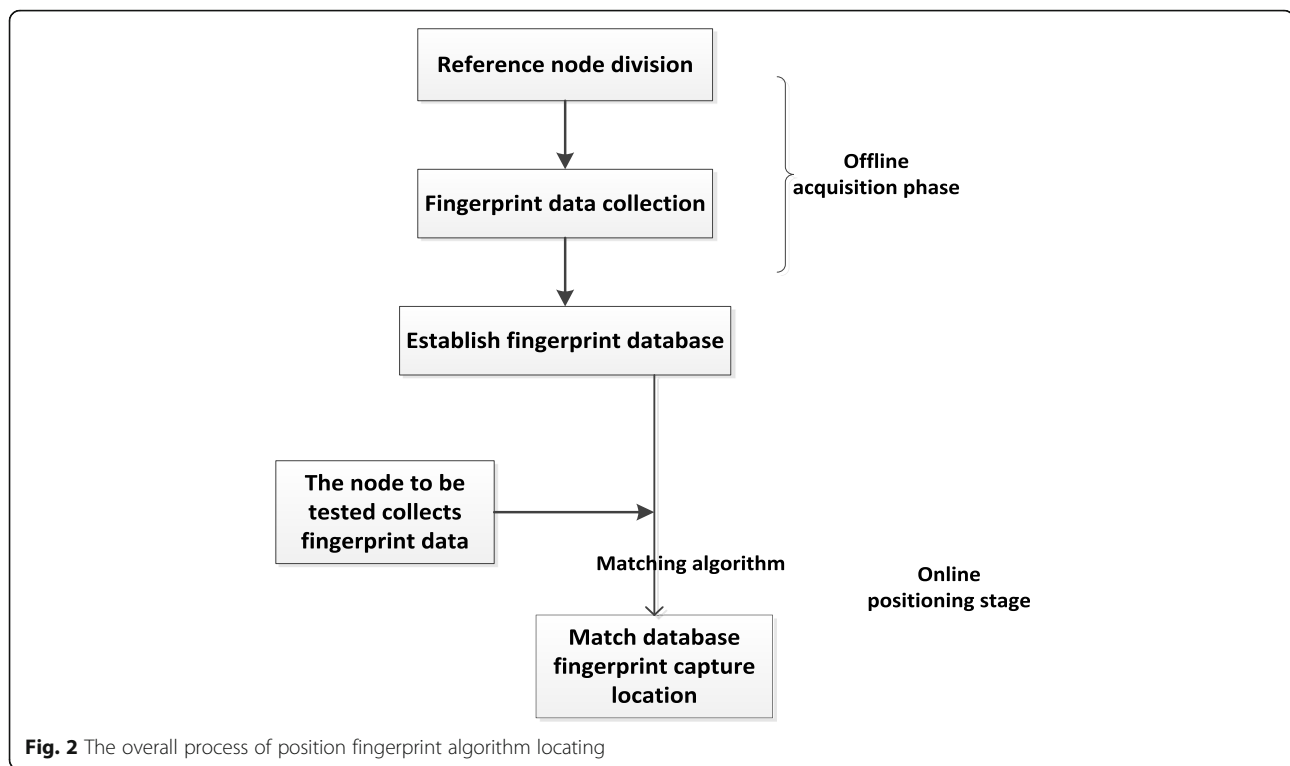
The offline acquisition stage generally needs to complete the following work steps (as shown in Fig. 3).

Step 1. The indoor positioning environment establishes a location metric and is divided into several regions. As shown in Fig. 3, the positioning area is divided into 16 small grids, and the position information of each grid is described by L_i , $i = 1, 2, \dots, 16$.

Step 2. Take signal strength information from different Wi-Fi APs. Figure 3 indicates that the location fingerprint information is being collected at the L_7 position. In general, there are two methods for obtaining AP signal strength information at a specific location. (1) Based on the terminal method, the signal strength information of a group of APs is collected one by one at the position of each reference point. (2) Based on the propagation model method, the signal strength information at each reference point is derived based on the position of the AP.

The terminal-based method achieves accurate signal strength and high positioning accuracy. Although it takes requires considerable time, and each new positioning area has to be re-acquired and built a location fingerprint database built, this method is adopted in this paper.

Step 3. Generate location feature fingerprint information at the reference point and store it in

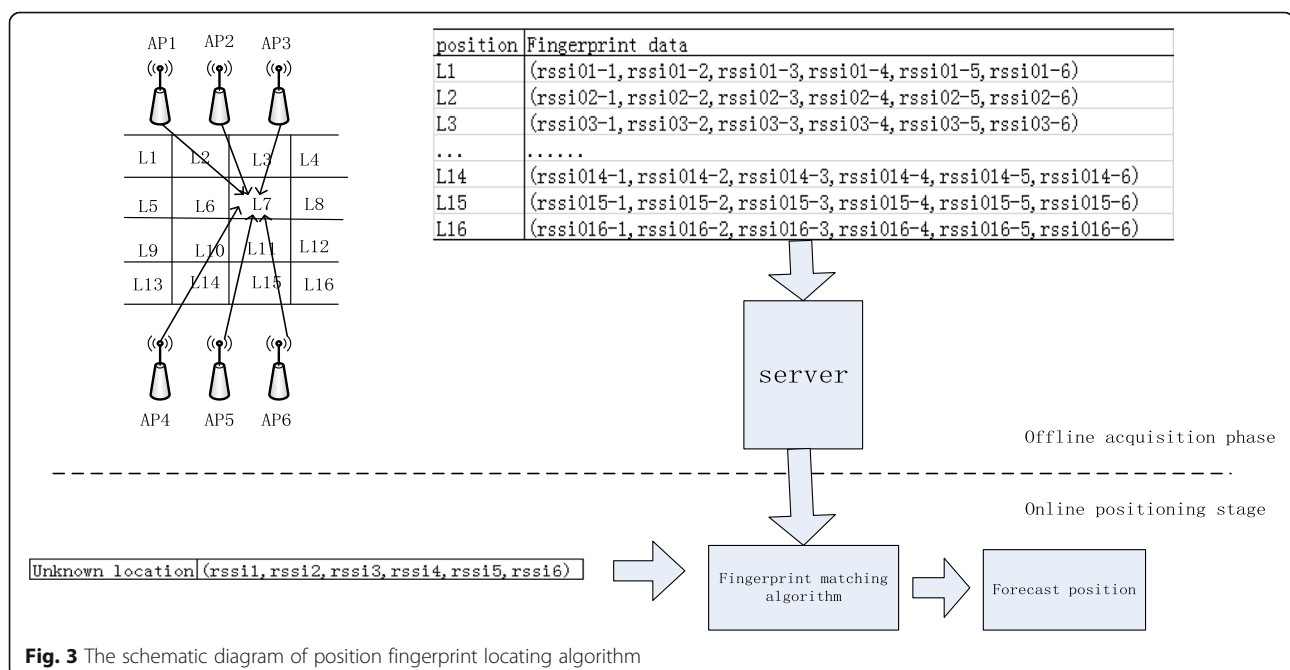


the location fingerprint database. In Fig. 3, the location fingerprint of the deterministic algorithm is generated; that is, the fingerprint at the location L_i is $(rss_{i1}-1, rss_{i1}-2, rss_{i1}-3, rss_{i1}-4, rss_{i1}-5, rss_{i1}-6)$, where rss_{ii} is the signal strength indicator; i is the number of the network; 1-6 are the numbers of the APs deployed in Fig. 3.

2.3.2 Online positioning stage

In the matching algorithm used in the positioning phase, the KNN algorithm used in this paper uses some metrics to distinguish the similarity between two different fingerprints, also called the dissimilarity.

1. Fingerprint dissimilarity



An informal definition of the degree of dissimilarity between two objects is a numerical measure of the degree of difference between the two objects. The more similar objects are, the lower their dissimilarity becomes. The distance is usually used as a synonym for dissimilarity, where distance is used to indicate the degree of dissimilarity between different nodes (such as the reference node and the point to be measured). Common methods are these, such as Euclidean distance, taxicab geometry, and the Minkowski distance.

2. Locating process

The matching algorithm during the locating process is divided into two types: deterministic positioning algorithm and probability-based positioning algorithm. This paper adopts a probability-based positioning algorithm.

The probability-based positioning algorithm stores the probability distribution model parameters of the collected signal strength information at different positions in the database (information such as histograms and Gaussian probability distributions are stored in the database), and Bayesian probability classification is used in the positioning phase to estimate the user's position. In the actual locating process, as shown in Fig. 3, the user-held mobile terminal collects a set of real-time location fingerprints at the unknown node position to be tested (rssii-1, rssii-2, rssii-3, rssii-4, rssii-5, rssii-6) and then uses a certain matching algorithm (e.g., KNN algorithm), applying the Euclidean distance in the fingerprint database. Measure the difference between the reference node and the node to be tested, and select the value of the K reference point with the smallest difference for test training. Too small or too large will affect accuracy.

A simplified version of the proposed algorithm can be described as follows:

Algorithm 1: k-Nearest Neighbor (KNN)

Input: Node to be tested S, $i=1, \text{distance}=0$

Begin

1. Calculate the size of the fingerprint database
2. Calculate the Euclidean distance between the point in the fingerprint database and the node to be tested
3. Sort by distance
4. Select K reference points with the smallest distance
5. Calculate the coordinates of the node to be tested
6. Average K reference points
7. Repeat steps 2-5 until all points to be measured are calculated.

End

Output: Position of the point to be measured

3 Improvement of the location fingerprint algorithm

The RSSI positioning method [20] is mainly divided into two categories: (1) based on theoretical or empirical propagation models, transforming propagation loss into

distance to achieve positioning; (2) based on pattern recognition, positioning by location fingerprint feature matching. Both are based on the intensity variation in the received signal, the difference being in the measurement index and whether the position of the reference node needs to be predicted [21].

3.1 Preprocessing of RSSI data

In the complex exhibition hall of a science and technology museum, the Wi-Fi signal is affected by various interference factors, and the signal attenuation is irregular. The signal strength of the same AP collected by the acquisition terminal at a fixed position is not fixed. In other words, the signal strength fluctuates within a certain range.

In this paper, the method of averaging multiple acquisition and measurement is used to smooth the calculated result to some extent, and thus, reduce the error caused by irregular fluctuations. However, in actual situations, even the fluctuation of different APs collected at the same position has different amplitude changes. Table 1 shows the Wi-Fi signal strength changes in different APs collected 2000 times at a fixed location.

The strength of the signal also changes according to the change in the position distance and is relatively stable within a certain distance range. After a certain range, it starts to increase. The Wi-Fi-based location fingerprint algorithm collects the signal strength feature values. To obtain more reliable location fingerprint feature data, the difference between each reference point and the node to be tested is more specific, which can reduce the positioning error, and improve the positioning accuracy in the positioning matching stage. This section uses the following methods to process the collected RSSI raw information data.

3.1.1 Truncated mean method

When the number of acquisitions is large enough, the arithmetic mean can represent the central trend of the sampling position, which is closer to the actual effective value of the position, and its calculation formula is Formula 1:

$$\bar{rssi} = \frac{\sum_{i=1}^N rssi_i}{N} = \frac{rssi_1 + rssi_2 + \cdots + rssi_N}{N} \quad (1)$$

N is the total number of acquisitions. Although the mean after a large number of samples is a useful measure to describe the RSSI data set, it is not the best method; there will be some irregular outliers, resulting in a larger error.

The truncated mean method can be used to eliminate the negative effect of outliers on the valid data. The core

Table 1 Signal amplitude change table for different APs with fixed positions

| AP name | Average value | Maximum value | Minimum value | Volatility |
|-----------|---------------|---------------|---------------|------------|
| TL-WR886N | − 42.58 dBm | − 35 dBm | − 64 dBm | 29 dBm |
| TL-WR742N | − 39.14 dBm | − 29 dBm | − 63 dBm | 29 dBm |
| STM-AUTO | − 57.14 dBm | − 53 dBm | − 66 dBm | 13 dBm |
| Mobile | − 47.29 dBm | − 37 dBm | − 64 dBm | 27 dBm |
| Complex | − 59.43 dBm | − 57 dBm | − 65 dBm | 8 dBm |
| STM | − 61.58 dBm | − 58 dBm | − 70 dBm | 12 dBm |

idea of this method is to discard the portion of the larger value end and the smaller value end, and then average the remaining data. For example, after sorting the collected data, discard 5% of the larger value end and 5% of the smaller value end, which the ratio must be controlled.

3.1.2 Screening mean method

Some researchers have proposed using the normal distribution model to preprocess the RSSI value. The model is based on the assumption that the RSSI value conforms to the normal distribution, that is, Formula 2 is satisfied.

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

$\mu = \frac{\sum_{i=1}^N x_i}{N}$ is the mean value and $\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{n-1}}$ is the standard deviation, where N is the number of acquisitions, and x_i is the RSSI value received during acquisition.

The theoretical probability value of the normal distribution in the interval is 0.954, and the theoretical probability value of the interval is 0.997.

Instead of using the normal distribution model directly, the normal distribution model and the trimmed tail estimation method can be comprehensively considered to establish a normal distribution model in the middle part where signal strength is relatively stable (which satisfies the premise), which can eliminate the outliers by the interval, to select the normal RSSI with higher probability.

Although there are many other methods for RSSI data preprocessing, for specific and complex indoor environments, it is necessary to consider the influence of various factors on Wi-Fi signals, and select different data processing methods according to local conditions to eliminate each collection work. The outlier data in the middle effectively improve the validity and reliability of the location fingerprint database, and ultimately improve the accuracy of Wi-Fi indoor positioning.

3.2 Algorithm improvement

The matching algorithm of the nearest neighbor classification algorithm in the Wi-Fi location fingerprint localization algorithm is based on the RSSI location fingerprint feature matching method to achieve position calculation. It compares the dissimilarity (or similarity) of the feature fingerprint information in the Wi-Fi feature fingerprint database with the Wi-Fi feature fingerprint information measured at the location of the unknown node to be determined, and thus determines the result of the location of the unknown node to be tested.

3.2.1 Improved algorithm based on distance weighting

The core idea of the distance-based weighted KNN algorithm is to assign different weighting parameters w_p according to the distance between the reference point and the point to be measured. A high weighting value is assigned to the node closer to the reference point and the reference point, while a low weighting value is assigned to the node far away from them. Therefore, the reference point closer to the point to be measured has a larger influence coefficient, and the farther reference point has a smaller influence coefficient, instead the influence coefficient in the conventional KNN algorithm is k (k is the latest number in the algorithm), and then the estimated position results are corrected to a certain extent, which improves the accuracy of positioning.

$$w_p = \frac{\frac{1}{d_{pj}}}{\sum_{i=1}^k \frac{1}{d_{pi}}} \quad (3)$$

Among them, $\frac{1}{d_{ij}}$ is the reciprocal of the Euclidean distance between the reference point i and the point j to be measured, and $\frac{1}{d_{pj}}$ is the reciprocal of the Euclidean distance from a reference point p of the k reference points to the point j to be measured.

From Formula 3, the coordinate calculation formula for the position of the point to be measured can be derived as:

$$(x, y) = \sum_{p=1}^k \frac{1}{\sum_{i=1}^k \frac{1}{d_{ij}}} (x_p, y_p) = \frac{1}{\sum_{i=1}^k \frac{1}{d_{1j}}} (x_1, y_1) + \dots + \frac{1}{\sum_{i=1}^k \frac{1}{d_{kj}}} (x_k, y_k) \quad (4)$$

According to Formula 4, the smaller the d_{pj} value is, the larger the reciprocal value is, and the larger the corresponding w_p value is, which increases the influence of the estimated position coordinate of the near reference point to the measured point.

Algorithm 2: Improved algorithm based on distance weighting

Input: Node to be tested S, $i=1, \text{distance}=0$

Begin

1. Calculate the size of the fingerprint database
2. Calculate the Euclidean distance between the point in the fingerprint database and the node to be tested
3. Sort by distance
4. Different weighting parameters are given according to the distance between the reference point and the point to be measured
5. Calculate the coordinates of the node to be tested according to the weighted distance
6. Average the calculated results
7. Repeat steps 2-6 until all points to be measured are calculated.

End

Output: Position of the point to be measured

Assuming that the quantity of data is expressed in n , the time complexity of the first, second, fourth, fifth and sixth steps in algorithm 2 is $O(n)$. The time complexity of step 3 is $O(n^2)$. Step 7 makes the time complexity of the whole algorithm $O(n^3)$.

3.2.2 Improved algorithm based on cosine similarity

Cosine similarity is a measure of similarity, assuming that x and y are two vectors to be compared, and the similarity function is shown in Formula 5.

$$\text{similarity}(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \cdot \|\vec{y}\|} \quad (5)$$

$\|\vec{x}\|$ represents the Euclidean norm of vector $\vec{x} = (x_1, x_2, \dots, x_p)$, which is the length of the vector $\sqrt{x_1^2 + x_2^2 + \dots + x_p^2}$. Similarly, $\|\vec{y}\|$ represents the Euclidean norm of the vector. The measure of similarity is the cosine between \vec{x} and \vec{y} . A cosine value of 0 means that the vectors are orthogonal and the angle is 90° . There is no similarity. The closer the cosine value is to 1, indicating a smaller angle, the greater the similarity in the vectors. The relationship between distance and cosine is shown in Fig. 4.

In the KNN matching algorithm, the cosine similarity can be selected as the position fingerprint between the reference point and the point to be measured. Calculate the cosine similarity between the point to be measured and each reference node. The following is an artificial pseudo code description for calculating the cosine similarity between two vectors.

Algorithm 3: Improved algorithm based on cosine similarity

Input: Node to be tested S, $i=1$

Begin

1. Calculate the size of the fingerprint database
2. Calculate the cosine similarity between the point in the fingerprint database and the node to be tested
3. Sort by cosine similarity
4. Calculate the coordinates of the node to be tested according to the cosine similarity
5. Average the calculated results
6. Repeat steps 2-5 until all points to be measured are calculated.

End

Output: Position of the point to be measured

The second step in algorithm 3 is to compute cosine similarity, and its time complexity is $O(n^2)$. The time complexity of this step is the highest in the whole algorithm. Step 6 makes the time complexity of the whole algorithm $O(n^3)$.

3.2.3 Improved algorithm for balancing joint metrics

The balanced joint metric algorithm is a trade-off balance algorithm for the above two algorithms. Assume that parameter is α . In the calculation of the nearest k reference points, distance weighting and cosine similarity are used as metrics to estimate the final position results of two points of the same point to be measured. Then, the results are weighted again, as shown in Formula 6.

$$(X_{\text{final}}, Y_{\text{final}}) = (1-\alpha) \cdot (X_{\text{weight}}, Y_{\text{weight}}) + \alpha (X_{\text{similarity}}, Y_{\text{similarity}}) \quad (6)$$

where $(X_{\text{final}}, Y_{\text{final}})$ is the final calculated position coordinate result, $(X_{\text{weight}}, Y_{\text{weight}})$ is the result of the position to be measured calculated by distance weighting as the metric, and $(X_{\text{similarity}}, Y_{\text{similarity}})$ is the cosine similarity as the metric calculated result of the position to be tested. When the value of α is equal to 0.5, it is equivalent to averaging the above two results; the specific values need to be dynamically adjusted according to different indoor environments. The following is the source code description of the simulation core part of cosine similarity.

Algorithm 4: Improved algorithm for balancing joint metrics

Input: Node to be tested S, $i=1$

Begin

1. Calculate the size of the fingerprint database
2. Calculate the cosine similarity between the point in the fingerprint database and the node to be tested
3. Sort by cosine similarity
4. Calculate your coordinates based on cosine similarity
5. Calculate the distance between the point in the fingerprint database and the node to be tested
6. Sort by distance
7. Different weighting parameters are given according to the distance between the reference point and the point to be measured
8. Calculate the coordinates of the node to be tested according to the weighted distance
9. The result of the calculation is weighted again to obtain the final result
10. Repeat steps 2-9 until all points to be measured are calculated.

End

Output: Position of the point to be measured

The highest time complexity in algorithm 4 is computational cosine similarity. Therefore, the time complexity of the algorithm and the time complexity of algorithm 3 are $O(n^3)$.

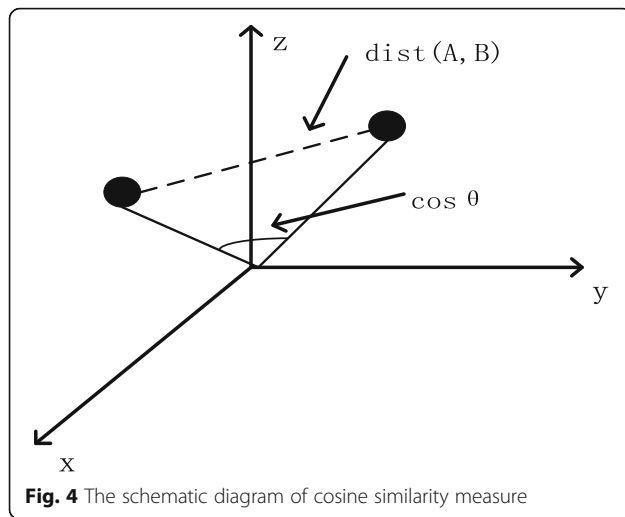


Fig. 4 The schematic diagram of cosine similarity measure

3.3 Summary of design ideas

In order to improve the positioning accuracy, we start from two aspects: data preprocessing and algorithm improvement.

- (1) Data preprocessing: Because Wi-Fi signal is affected by various interference factors and its attenuation is irregular, the original data need to be preprocessed. The main methods are truncated mean method and screening mean method. The core idea of truncated mean method is to discard the portion of the larger value end and the smaller value end, and then average the remaining data. The screening mean method establishes a normal distribution model in the middle part of the relatively stable signal strength, and eliminates abnormal values in this way. Truncated mean method and screening mean method make the location fingerprint data more effective and reliable, and provide a guarantee for improving the location accuracy.
- (2) Algorithm improvement: The core idea of the distance-based weighted KNN algorithm is to assign different weighting parameters according to the distance between the reference point and the point to be measured, to assign higher weighting values to the nodes whose distance between the reference point and the point to be measured, and to assign lower weighting values to the nodes whose distance is closer, instead of the traditional KNN algorithm where the influence coefficients are all $1/k$ (k is the number of the nearest neighbors in the algorithm), so as to estimate the location. The results are corrected to improve the positioning accuracy. The idea of the improved algorithm based on cosine similarity is that the cosine value is closer to 1, which indicates that the smaller the angle is, the greater the similarity between vectors. Choosing

cosine similarity as the measurement index of fingerprint feature vector between reference point and test point can improve the positioning accuracy. Balancing joint metrics algorithm is a balancing algorithm which combines the above two algorithms. The main idea is to use distance weighting as the measurement index to calculate the result of the position to be measured, cosine similarity as the measurement index to calculate the result of the position to be measured, and then to average the two results, so as to improve the positioning accuracy.

The experiment in Section 4 shows the accuracy of the design idea.

4 Simulation experiment and analysis

4.1 Fingerprint positioning environment

The $10.5 \times 18 \text{ m}^2$ areas of a science and technology museum were selected as the indoor positioning experimental environment in Fig. 5. First, the test area is modeled, and the upper left corner of the area is used as the coordinate origin to establish a two-dimensional coordinate system.

4.2 Location fingerprint collection

In the offline sampling measurement phase, an Android terminal is used as a fingerprint information device for collecting Wi-Fi signals, and the mobile phone model is MI-A. In the acquisition stage, set a scan interval of 5 seconds, collect 66 sets of RSSI raw data for all reference points, save them to the local file in txt format, upload them to the server through Wi-Fi (or a mobile communication network), and build on the server side. After the fingerprint data are built on the server side, a mapping relationship between a coordinate point (x, y) and a strong vector of Wi-Fi signals is established to form an offline fingerprint database. In the actual acquisition, there are 6APs with stable Wi-Fi signals for reference, and the others are filtered. Then, at each reference point, 66 sets of fingerprint data of the location feature fingerprint and the point coordinates are generated to complete the

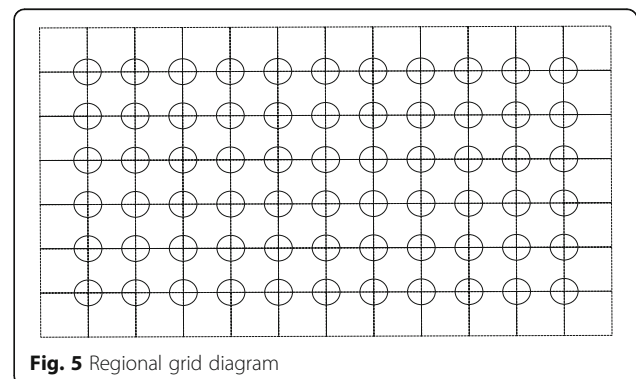


Fig. 5 Regional grid diagram

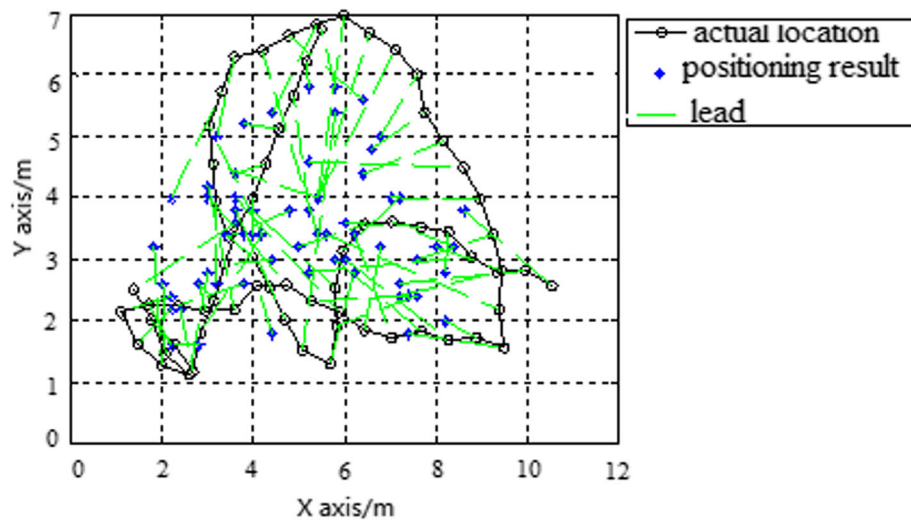


Fig. 6 The result of positioning analysts based on traditional KNN

construction of the fingerprint database. The mapping format is as follows:

$\langle \text{Loc}_x, \text{Loc}_y, \text{Rssi-1}, \text{Rssi-2}, \text{Rssi-3}, \text{Rssi-4}, \text{Rssi-5}, \text{Rssi-6} \rangle$.

4.3 Locating result simulation

This article uses MATLAB R2013b as the simulation environment, and under the premise that 66 reference point offline location fingerprint databases have been constructed, 70 sets of measured points of measured position fingerprint data are selected as analysis indicators. Different improved algorithms are selected as the positioning result analysis.

- (1) Positioning analysis based on traditional KNN algorithm

The positioning result of the traditional KNN algorithm using Euclidean distance as a metric is shown in Fig. 6. Is the comparison between the actual position generated and the positioning estimation result with the minimum deviation of average positioning distance. The minimum deviation of average positioning distance is 1.7105 m, and $k = 5$.

- (2) Positioning analysis based on distance weighting improvement

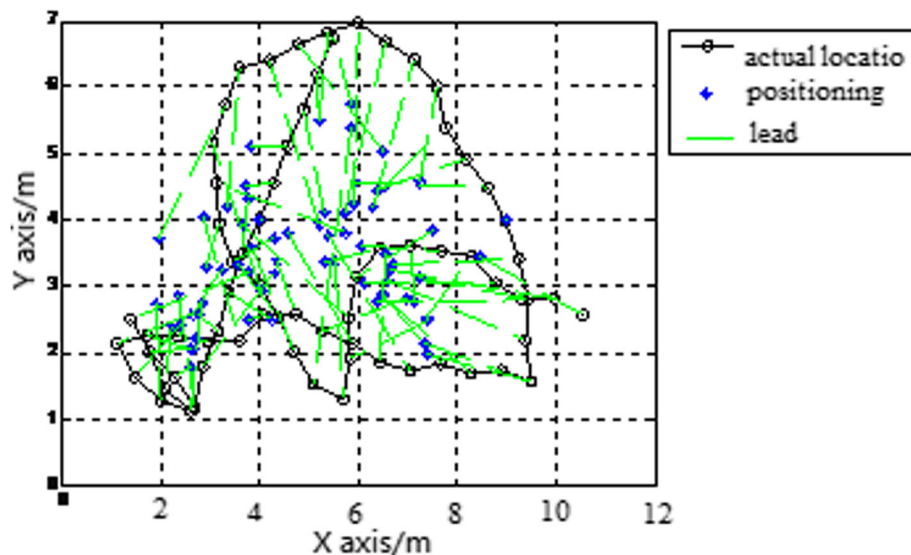


Fig. 7 The result of positioning analysts based on distance weighting KNN

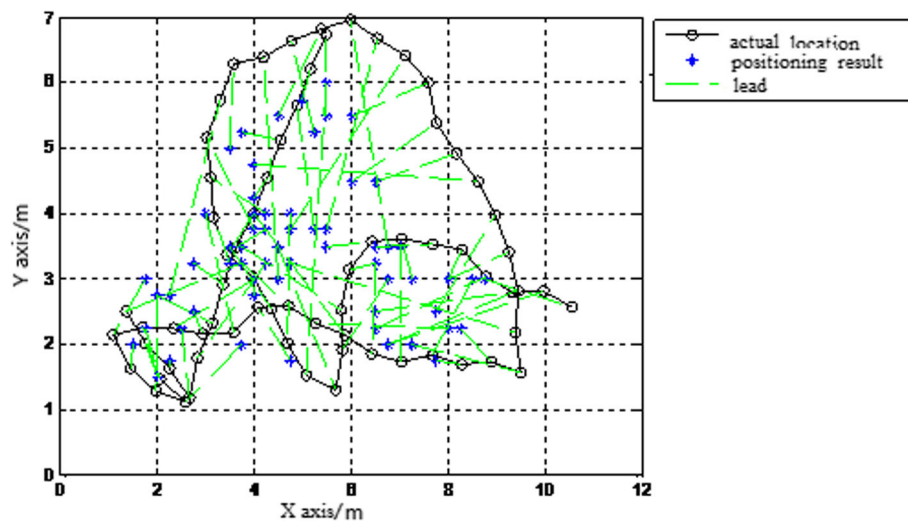


Fig. 8 The result of positioning analysts based on cosine similarity

The result of positioning analysis based on KNN algorithms of distance weighting improvement is shown in Fig. 7. In the figure, the average positioning deviation of the 70 groups of nodes to be tested is the smallest, and the deviation is 1.7105 m, and $k = 7$. It can be found that the positioning accuracy is improved compared with the traditional Euclidean distance positioning algorithm.

(3) Improved positioning results based on cosine similarity

When the cosine similarity is selected as the metric to improve the algorithm, it is found that while k is taken as 4, the average positioning error of the node to be tested is small, which is 1.6549 m. The comparison of the positioning results at this time is shown in Fig. 8. The positioning accuracy is improved, comparing with the traditional KNN algorithm based on Euclidean distance.

(4) The result of positioning analysts based on balanced joint metric improvement algorithm

The situation of balancing joint metrics is slightly complicated. First, assuming that a takes a value of 0.5, then separately calculate the effect of k taking a values from 1 to 10 on the positioning accuracy. The result is shown in Table 2.

Obviously, the positioning accuracy is higher when $k = 4$. k takes the value of 4, and the influence of the joint average metric improvement algorithm on the positioning deviation is shown in Table 3.

The joint metric algorithm model shows that when $a = 0$, it is an algorithm based on distance weighting. When $a = 1$, it is an algorithm based on similarity

metric. Combining Tables 2 and 3, select $k = 4$, $a = 0.7$ for the positioning result analysis, as shown in Fig. 9.

(5) Comparison of positioning deviation of three improved algorithms

After completing the analysis of the positioning results of different improved algorithms, different k values are selected here. The traditional positioning algorithm based on Euclidean distance and the positioning deviation between the improved algorithms of three different strategies are comprehensively compared and analyzed. In Fig. 10, it is found that the positioning accuracy improved by the improved algorithm of three different strategies compared with the traditional algorithm. The improved algorithm based on distance weighting and the balanced joint metric algorithm are smoother and more stable, and the joint balance metric algorithm performs better than another, which preliminarily verifies the effectiveness and superiority of the joint balance metric algorithm.

(6) Comparisons of algorithms

Table 2 The influence of the value of k on the positioning deviation of joint metrics

| The value of K | Positioning deviation (m) | The value of K | Positioning deviation (m) |
|----------------|---------------------------|----------------|---------------------------|
| 1 | 2.1264 | 6 | 1.6915 |
| 2 | 1.9458 | 7 | 1.6844 |
| 3 | 1.6473 | 8 | 1.7189 |
| 4 | 1.6466 | 9 | 1.7269 |
| 5 | 1.6661 | 10 | 1.7917 |

Table 3 The influence of the value of α on the positioning deviation of joint metrics when $k = 4$

| Value of α | Positioning deviation (m) | Value of α | Positioning deviation (m) |
|-------------------|---------------------------|-------------------|---------------------------|
| 0 | 1.7145 | 0.5 | 1.6466 |
| 0.1 | 1.6952 | 0.6 | 1.6416 |
| 0.2 | 1.6791 | 0.7 | 1.6396 |
| 0.3 | 1.6649 | 0.8 | 1.6411 |
| 0.4 | 1.6541 | 0.9 | 1.6462 |

We performed experiments in various regions, such as $20\text{ m} \times 20\text{ m}$ and $15\text{ m} \times 24\text{ m}$. Although the accuracy of the joint balance metric improvement algorithm is not the best, the algorithm is simple and the time complexity is only $O(n^3)$.

The accuracy of the neural network algorithm is higher [18]. Global time complexity of convolutional Neural Networks is: $\text{times} \sim O(\sum_{l=1}^D M_l^2 K_l^2 C_{l-1} C_l)$. Among them: D is the number of convolution layers of the neural network; L is the first convolution layer of the neural network; M is the edge length of the output characteristic graph of each convolution core; K is the edge length of each convolution core; C_l is the number of convolution kernels of the l th layer.

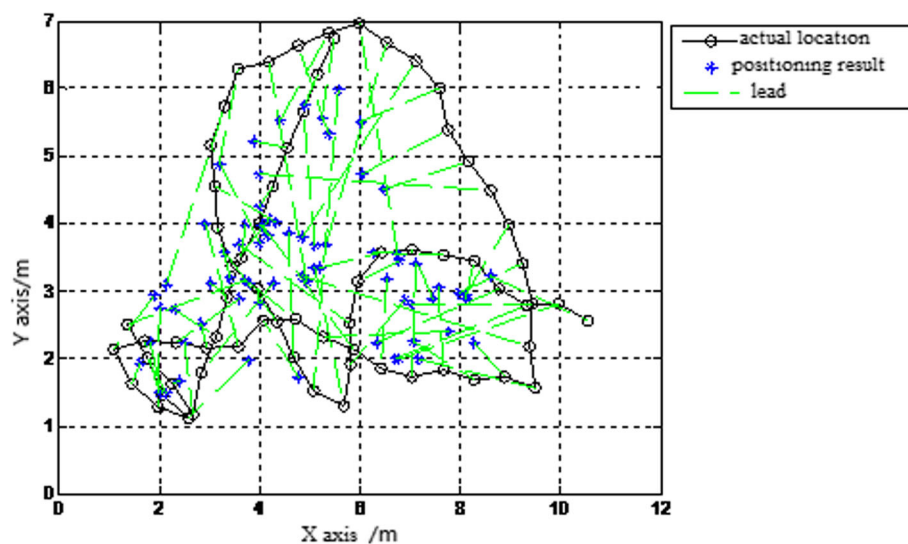
It can be seen that the time complexity of the algorithm we designed is lower. This algorithm can be applied to general science and technology museums.

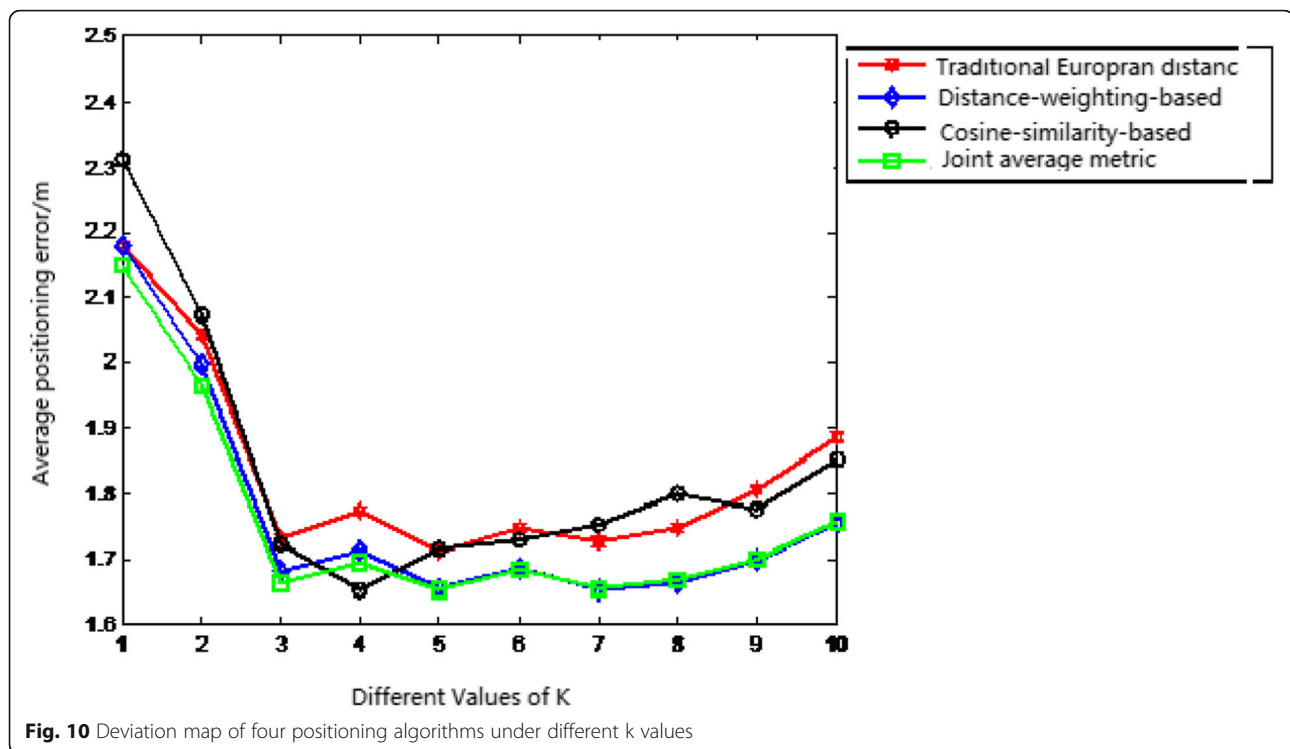
5 Conclusions

In recent years, with the development of science and technology and the progress of society, the science and

technology museums have become increasingly popular. The main concern of the manager is to improve the quality of the science and technology museum's exhibition hall and to understand its popularity. This article uses a phone as the mobile terminal, collecting visitors' trajectories through indoor positioning technology, which is helpful for the scientific management of the science and technology museum. Today's indoor segment distance positioning is still in the stage of exploration and research, and there are no large-scale deployments applied in China. Due to the high popularity of mobile terminals and extensive coverage of Wi-Fi networks, this article chooses Wi-Fi technology based on mobile terminals for the positioning of users of the science and technology museum, improves the location fingerprint algorithm, and illustrates its effectiveness in experiments.

Although this paper proposes some methods for processing the RSSI raw data of the location fingerprint algorithm, provides three improved schemes for the KNN matching algorithm and improves the validity of fingerprint data and optimization of algorithm accuracy, it

**Fig. 9** Positioning result of joint balance metric improvement algorithm



only considers the location fingerprint information of two-dimensional space in the simulation experiment to simplify the research work. In fact, the real environment in the room is three dimensional, and the fingerprint information at different heights will be different even in the same position, which is ignored in this article. Therefore, to obtain higher positioning accuracy, it is necessary to consider the construction of the location fingerprint database in three-dimensional space.

Abbreviations

AHLs: Ad hoc localization system; AOA: Angle of arrival; AP: Wireless access point; APs: Approximately 6 wireless access points; BSSID: Basic service set; KNN: k-nearest neighbor; LAN: Local area network; MIT: Massachusetts Institute of Technology; PS: Photoshop; RADAR: Radio detection and ranging; RFID: Radio frequency identification; RSSI: Received signal strength indication; RSSP: Received signal strength; SI: Received; SSID: Service set identifier; TDOA: Time difference of arrival; TOA: Time of arrival; UCLA: University of California, Los Angeles; UWB: Ultra wide band; Wi-Fi: Wireless-fidelity

Acknowledgements

The authors thank the anonymous reviewers and editors for their efforts in valuable comments and suggestions.

Authors' contributions

QY proposes the algorithm design and data analysis. SZ proposes innovation ideas. ML proposes data analysis. YZ carried out experiments. All authors read and approved the final manuscript.

Funding

This paper is supported by the national science and technology support Program. Program number is 2015BAK33B00.

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Competing interests

The authors declare that they have no competing interests.

Received: 27 May 2019 Accepted: 11 November 2019

Published online: 17 December 2019

References

1. C. Jin, D. Qiu, Research on indoor positioning technology based on WiFi signal [J]. *Bull Survey Mapp* **2017**(05), 21–25 (2017)
2. Cheng Y C, Chawathe Y, Lamarca A, Accuracy characterization for metropolitan-scale Wifi localization. 2005 Proceedings of the 3rd International Conference on Mobile System, Applications and Services (MobiSys, Washington, 2005), pp. 233–245. <https://dl.acm.org/citation.cfm?doi=10.67170.1067195>. Accessed Oct 2017.
3. L.M. Ni, Y. Liu, Y.C. Lau, A.P. Patil, LANDMARC: indoor location sensing using active RFID [J]. *Wirel Netw* **10**(6), 701–710 (2004)
4. Lee J S, Su Y W, Shen C C, A comparative study of wireless protocols: Bluetooth, UWB, ZigBee, and Wi-Fi. IEECON 2007-33rd Annual Conference of the IEEE Industrial Electronics Society IEEE. (Taiwan, 2007), pp. 46–51. <https://ieeexplore.ieee.org/document/4460126>. Accessed Oct 2017.
5. W. Rong, Brief discussion on the popular short-distance wireless communication technology nowadays. *Youth Era* (14), 199–200 (2015, 2015)
6. Wei-qing Huang, Chang Ding, Si-ye Wang, Junyu Lin, Shao-yi Zhu, Yue Cui, Design and Realization of an Indoor Positioning Algorithm Based on Differential Positioning Method. *International Conference on Wireless Algorithms, Systems, and Applications*. (WASA, Guilin, 2017), pp. 546–558. https://link.springer.com/chapter/10.1007/978-3-319-60033-8_47. Accessed Oct 2018.
7. Yucel, Edizkan, Ozkir, Yazici, Development of indoor positioning system with ultrasonic and infrared signals. *International Symposium on Innovations in Intelligent Systems & Applications*. (Trabzon, 2012). <https://ieeexplore.ieee.org/document/6246983>. Accessed May 2018.

8. Alessio De Angelis, Antonio Moschitta, Paolo Carbone, Massimo Calderini, Stefano Neri, Renato Borgna, Manuele Peppucci, Design and characterization of an ultrasonic indoor positioning technique. *Instrumentation and Measurement Technology Conference (I2MTC) Proceedings 2014 IEEE International*. (Montevideo, 2014), pp. 1623-1628. <https://ieeexplore.ieee.org/document/6861020>. Accessed May 2018.
9. Lingling Zhu, Aolei Yang, Dingbing Wu, Li Liu, Survey of Indoor Positioning Technologies and Systems. *2014 International Conference on LifeSystem Modeling and Simulation*. (Shanghai, 2014), pp. 400-409. https://link.springer.com/chapter/10.1007/978-3-662-45283-7_41. Accessed Oct 2017.
10. Want R, Hopp A, Falcão V, Gibbons J, The active badge location system. *Acm Transactions on information Systems*. 10(1), 91-102(1992)
11. Bahl P, Padmanabhan V N, RADAR: an in-building RF-based user location and tracking system. *Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*. (IEEE, Tel Aviv, 2000), pp. 775-784. <https://ieeexplore.ieee.org/document/832252>. Accessed Mar 2018.
12. Woodman O J, Harle R K, Concurrent scheduling in the Active Bat location system. *IEEE International Conference on Pervasive Computing and Communications Workshops*. (Mannheim, 2010), pp. 431-437. <https://ieeexplore.ieee.org/document/5470631>. Accessed Mar 2018.
13. Priyantha N, Chakraborty A, Balakrishnan H, The Cricket location-support system. *Proceedings of the 6th annual international conference on Mobile computing and networking*. (Boston, 2010), pp. 32-43.
14. Oliveira H A B F, Boukerche A, Nakamura E F, Localization in time and space for wireless sensor networks: an efficient and lightweight algorithm. *Performance Evaluation*. 66(35), 209-222 (2009)
15. D.J. Coleman, A. Rajabifard, K.W. Kolodziej, Expanding the SDI environment: comparing current spatial data infrastructure with emerging indoor location-based services. *Int J Digital Earth* 9(6), 629-647 (2016)
16. Hanhui Yue, Xiao Zheng, Juan Wang, Li Zhu, Chunyan Zeng, Cong Liu, Meng Liu, Research and Implementation of Indoor Positioning Algorithm for Personnel Based on Deep Learning. *International Conference on Emerging Internetworking, Data & Web Technologies*. (Tirana, 2018), pp. 782-791. <https://link.springer.com/content/pdf/10.1007%2F978-3-319-75928-9.pdf>. Accessed Feb 2019.
17. He Z, Xu M, Guo A, Advanced Computational Methods in Life System Modeling and Simulation. Multi-channel feature for pedestrian detection (Nanjing, 2017), pp. 472-480. https://link.springer.com/chapter/10.1007/978-981-10-6370-1_47. Accessed Oct 2018.
18. J. Luo, H. Gao, I. Gondal, Deep belief networks for fingerprinting indoor localization using ultrawideband technology [J]. *Int J Distributed Sensor Netwo* (2016)
19. Aguilar, W.G., Luna, M.A., Moya, J.F., Abad, V., Ruiz, H., Parra, H., Lopez, W, Cascade classifiers and saliency maps based people detection. *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*. (Otranto, 2017), pp. 501-510.
20. Jaemin Hong, KyuJin Kim, ChongGun Kim, Comparison of Indoor Positioning System Using Wi-Fi and UWB. *Asian Conference on Intelligent Information and Database Systems*. (Dong Hoi City, 2018), pp. 623-632. <https://link.springer.com/content/pdf/10.1007%2F978-3-319-75417-8.pdf>. Accessed Oct 2018.
21. Jianguo Yu, Zhian Deng, Xin Liu, Juan Chen, Zhenyu Na, WLAN Indoor Positioning Based on DLDA Feature Extraction Algorithm. *International Conference in Communications, Signal Processing, and Systems*. (Harbin, 2017), pp. 2779-2787. https://link.springer.com/chapter/10.1007/978-981-10-6571-2_336. Accessed Feb 2019.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)