

RESEARCH

Open Access



A method of fingerprint indoor localization based on received signal strength difference by using compressive sensing

Xiao-min Yu^{1,2}, Hui-qiang Wang^{1*} and Jin-qiu Wu³

Abstract

With the development of network technology, WLAN-based indoor localization plays an increasingly important role. Most current localization methods are based on the comparison between the received signal strength indication (RSSI) and the RSS in the database, whose nearest reference point is the location point. However, since a uniform standard for measuring components of smartphones has not yet been established, the Wi-Fi chipsets on different smartphones may have different sensitivity levels to different Wi-Fi access points (APs) and channels. Even for the same signal, RSSI values obtained by different terminals at the same time and the same location may be different. Therefore, the impact of terminal heterogeneity on localization accuracy can be overlooked. To address this issue, a fusion method based on received signal strength difference and compressive sensing (RSSD-CS) is proposed in this paper, which can reduce the influence caused by the terminal heterogeneity. Besides, a fingerprint database is reconstructed from the existing reference point data. Experiments show that the proposed RSSD-CS algorithm can achieve high localization accuracy in indoor localization, and the accuracy is enhanced by 20.5% and 15.6% compared to SSD and CS algorithm.

Keywords: WLAN, Wi-Fi, Indoor localization, Received signal strength difference, Compressive sensing

1 Introduction

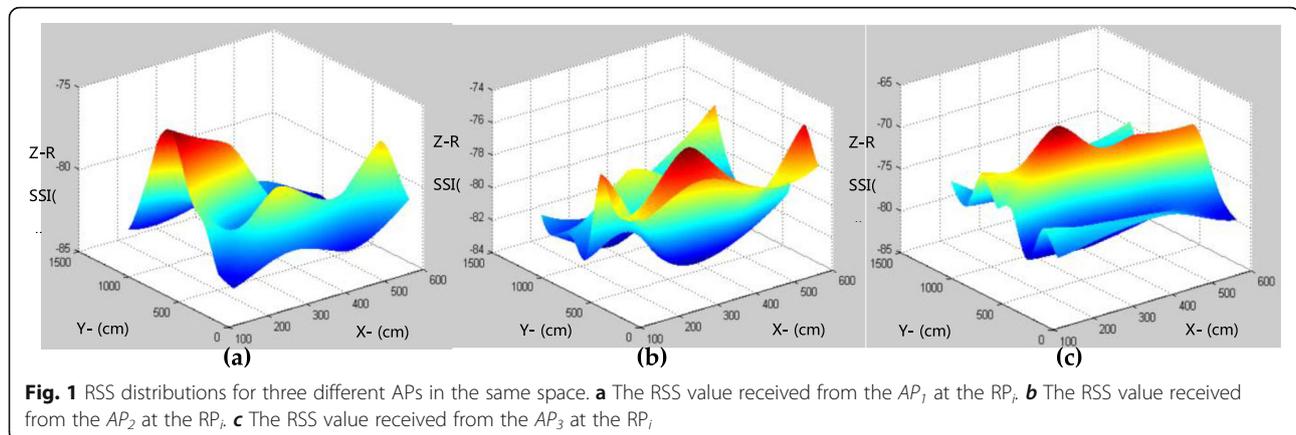
The rapid development of smartphones has made it a carrier of location-based service (LBS), such as indoor localization, navigation, and tracking. In recent years, received signal strength (RSS) fingerprint-based Wi-Fi positioning methods have attracted the attention of many researchers because RSS can be easily obtained by a Wi-Fi-integrated mobile device without any additional hardware [1–5]. The performance of the fingerprint-based methods depends on the number of the reference points (RPs) in unit space. However, as RSS measurement is time-consuming and laborious, an increase in the number of RPs will increase positioning costs.

Figure 1 shows the RSS distribution of three different access points (APs) in the same space. It can be seen that the RSS distribution of different APs in the same space is different because of the differences of their place, so there is a corresponding relationship between the location and the RSS values, just like the fingerprints of persons. It is called “location fingerprint.”

Fingerprint-based positioning requires first establishing a fingerprint database and then performing online matching in two stages of rough matching and precise matching in this paper. Wireless signals depend on the propagation environment, and the multi-path characteristics of different channels are different. During the transmission of wireless signals, unique signals related to the transmission environment are produced through reflection, refraction, and scattering. The multi-path feature is referred to as the “location fingerprint” [6]. However, since the radio frequency (RF) signals can vary

* Correspondence: wanghuiqiang@hrbeu.edu.cn

¹College of Computer Science and Technology, Harbin Engineering University, Harbin, People's Republic of China
Full list of author information is available at the end of the article



over time and space due to obstacles and multi-path fading, this may degrade performance of using RSS fingerprints for positioning. It is therefore difficult to achieve good accuracy in most applications.

To achieve better indoor localization results, there are many fusion approaches to enhance the positioning accuracy [2, 7, 8]. However, if different types of sampling terminals are employed in offline stages and online stages, the positioning accuracy will be reduced directly [9]. Such device heterogeneity will affect the performance of indoor localization [10]. Therefore, the key of indoor localization technology is to design an accurate real-time indoor localization system without any hardware installation and modification, which can be readily deployed to the mobile devices.

A method by the existing Wi-Fi infrastructure, without the need of knowing the specific location of each AP, is proposed in this paper. The compressive sensing technology is used for reconstructing the position fingerprints, and the received signal strength difference (RSSD) is used for replacing the traditional RSS. The purpose is to improve the positioning accuracy without increasing the hardware investment, no matter which devices the users use.

The main contributions of this paper are as follows:

- To present a method of indoor positioning through the fusion of RSSD and compressive sensing (CS) without any hardware installation and modification. Replace the traditional RSS with RSSD to realize localization, which can undermine the influence of device heterogeneity on localization accuracy to some extent. At the same time, the compressive sensing algorithm is presented for location matching. The existing reference point data are used for restoring the complete fingerprint database so that every point in space has a corresponding fingerprint and the localization accuracy can be improved to a certain extent.

- To present an improved fuzzy clustering method (IFCM) algorithm, unlike the traditional fuzzy clustering method (FCM) algorithm, it is not necessary to determine the clustering center in advance. It takes all sample points as the candidate clustering centers and gives each sample point a corresponding real number, which is called the biased parameter. The larger the deviation parameter is, the more likely to be determined as the clustering center.

The rest of this paper is organized as follows: Section 2 describes a summary of related work. Next, Section 3 describes the overall system architecture and RSSD-CD methods. The experimental results and the related discussion are presented in Section 4. Finally, Section 5 provides the conclusions of this paper.

2 Related work

Fingerprint-based location technology does not require prior knowledge of infrastructure location, nor does it need a propagation model. In general, fingerprint-based methods map RSS values to physical locations through pattern recognition technology [11]. This kind of RSS-localization mapping is obtained from the site survey, and it can be located by using different statistical modeling or machine learning algorithm to compare data in online measurement and pre-existing database [12, 13]. K -nearest neighbor (KNN) [14], multi-layer perceptron [15], neural network [16, 17], maximum likelihood [18, 19], support vector machine [20], nuclear methods [21–23], etc. can be used in indoor localization. Nuclear-based machine learning technology can also be applied to WLAN localization [21]. Indoor signal propagation environment, however, is quite complex. The RSS of a given location usually changes over time due to the reason of interference, reflection, refraction, multi-path fading, direction of equipment, hardware changes, temperature, humidity, and even attack [24–29]. Additional radio bandwidth will also change RSS in a bandwidth-constrained system [30, 31].

Such noise causes the mismatch between the online measuring data and the offline recorded data in the location fingerprint system. The previously stored fingerprinting map no longer reflects the statistical features of the current RSS, so the performance of the system decreases. In an actual indoor environment, this problem is inevitable. At the same time, the Wi-Fi chipsets on different smartphones may be sensitive to different Wi-Fi APs and channels. Therefore, the signal values and length of signal vector [32] are different. The operating systems of smartphones may support different APs. The detection rate and number of APs sensed can be altered [33]. Therefore, even the same device with different operating systems may still have heterogeneity.

In literature [34, 35], the authors proposed that RSS for different devices follow a linear model for the same signals. The signals of two devices are given, and the linear calibration model is obtained by regression. Offline training is required before the exact mapping relationships. However, manual data acquisition and calibration are inconvenient for deployment. To reduce such labor-intensive work, some crowd-sourced algorithms [36] can be implemented to provide efficient signal data for calibration. To reduce the workload of offline calibration, online calibration is adopted in literature [32, 37]. Hosain et al. have proposed a simple signal strength difference (SSD) algorithm based on the signal and antenna gain model. The constant factor of the antenna gain is derived. A similar idea is employed in the received signal strength indication (RSSI) ratio [37]. An AP measurement value is selected [37], and each AP signal is divided by this constant value. These methods are easy to realize on the online calibration of the signal vector.

Numerous researchers improve the accuracy of the indoor localization system from different aspects. Wang et al. [8] have used curve fitting (CF) technology to determine RSS-distance relation in an indoor environment, instead of using a logarithmic distance model and a series of linearly independent functions to build a generic RSS-distance model. The advantage of this approach is that the data can be extracted from a small number of RSS measurements. By dividing space and curve fitting in each partition, a more accurate RSS-distance relationship can be obtained for each AP. First, a subarea which a mobile device belongs to is determined, and then the sum of distance errors can be minimized to realize localization. However, the drawback of this scheme is that it only considers the regular layout of the room and not considers the diversity of terminals.

Fang et al. [38] have matched the RSS value of the mobile device and the fingerprint database by using the compressive sensing principle. Compressive sensing technology can be applied to compensate for the scarcity compressed signal or noise. This method uses a small

amount of fingerprint sampling to reconstruct a complete fingerprinting map, it can reduce the workload of building a fingerprint database and also can improve the accuracy of positioning, but this method still does not take the diversity of terminals into account and has a higher computational complexity.

Wen et al. [39] have proposed a new algorithm to adjust the RSS fingerprint database by using the feedback information about the surrounding environment. The system uses different, updated measurements and offline fingerprint points according to the time and spatial intensity of the location. Because the adaptive radius is applied as the main factor of affecting positioning accuracy, this method can solve the problem of signal occlusion which were caused by moving obstacles in simulation and actual scenes.

Zheng et al. [40] have used a low-tubal-rank tensor to model Wi-Fi fingerprints of all reference points (RPs) and proposed an adaptive sampling scheme via approximate volume sampling to improve reconstruction accuracy of radio map with reduced expenditure. Meanwhile, they have provided a mathematical theory to analyze and derive the performance bounds of the proposed method in terms of sample complexity and reconstruction error.

References [41–43] are based on outdoor positioning. Wu et al. [41] proposed a mobile positioning system and a mobile positioning method based on recurrent neural networks to analyze the RSSIs from heterogeneous networks, which include cellular networks and Wi-Fi networks. The network signals from heterogeneous networks can be analyzed to improve the accuracies of the estimation of locations. Cheng et al. [42] proposed an intelligent positional approach for high-speed trains based on ant colony optimization and machine learning algorithms. The proposed methods can enhance the real-time performance in the online updating process on the premise of reducing the positioning error. Chen et al. [43] proposed a two-stage estimation algorithm based on variable projection method for GPS positioning. The proposed method can effectively mitigate multi-path interference.

Many researchers have used different methods to improve indoor localization performance. However, they do not consider the influence of both equipment heterogeneity and non-reference point factors on positioning accuracy. Compared to these works, our received signal strength difference and compressive sensing (RSSD-CS) method can restore complete signal space with limited RSS measurement. It can reduce noise interference and remove outliers to some extent. No matter what type of terminals are being used, a satisfactory positioning effect can be obtained.

3 Methods

Although locational fingerprinting technologies based on Wi-Fi have relatively high localization accuracy, collected RSS samples could be affected by indoor environmental

factors, such as multi-path, shielding, and person’s movement. RSS values are likely to be different even at the same device, same location, and different time, which could affect localization accuracy significantly. Due to the differences in the type of chipsets in the devices, the RSS values measured by different devices at the same time, in the same location, are also different. It is impossible to establish a fingerprint database for each device in the database. Therefore, the error caused by the heterogeneity of devices cannot be ignored.

As can be seen from Fig. 2, the sampling of RSS values at a different time in the same location is constantly changing. The estimation of location only by absolute RSS value will produce a large error. It is necessary to select the mean value within continuous time sampling as the measured RSS values for calculation both online and offline states. Doing so can reduce the adverse impact of environmental factors, but it is impossible to eliminate it fundamentally.

The indoor propagation model of the wireless signal is [44]:

$$\begin{aligned}
 P(d)|_{\text{dBm}} &= P(d_0)|_{\text{dBm}} - 10\beta \lg\left(\frac{d}{d_0}\right) \\
 &+ X_{\text{dB}} = 10 \lg\left(\frac{P_{\text{AP}}G_{\text{AP}}G_{\text{MT}}\lambda^2}{16\pi^2d_0^2L}\right) - 10\beta \lg\left(\frac{d}{d_0}\right) + X_{\text{dB}}
 \end{aligned}
 \tag{1}$$

where P_{AP} is the transmission power of AP, G_{AP} is the antenna gain of AP, G_{MT} is the antenna gain of mobile terminal (MT), L is system loss factor, λ is the wavelength of the wireless signal. Due to different Wi-Fi network interface controllers (NICs) embedded in smartphones, the antenna gain (G_{AP} and G_{MT}) may be different. This leads to RSS samples at the same time, and the same location by different devices are also different.

From formula (1), it can be seen that the RSS value at a distance d depends on the hardware parameter of MT. So,

$$P(d_1)|_{\text{dBm}} = 10 \lg\left(\frac{P_{\text{AP1}}G_{\text{AP1}}G_{\text{MT}}\lambda_{\text{AP1}}^2}{16\pi^2d_0^2L_1}\right) - 10\beta_1 \lg\left(\frac{d_1}{d_0}\right) + [X_1]_{\text{dB}}
 \tag{2}$$

$$P(d_2)|_{\text{dBm}} = 10 \lg\left(\frac{P_{\text{AP2}}G_{\text{AP2}}G_{\text{MT}}\lambda_{\text{AP2}}^2}{16\pi^2d_0^2L_2}\right) - 10\beta_2 \lg\left(\frac{d_2}{d_0}\right) + [X_2]_{\text{dB}}
 \tag{3}$$

Formula (2) minus formula (3) is:

$$\begin{aligned}
 P(d_1)|_{\text{dBm}} - P(d_2)|_{\text{dBm}} &= 10 \lg\left(\frac{P_{\text{AP1}}G_{\text{AP1}}\lambda_{\text{AP1}}^2L_2}{P_{\text{AP2}}G_{\text{AP2}}\lambda_{\text{AP2}}^2L_1}\right) \\
 &- 10\beta_1 \lg\left(\frac{d_1}{d_0}\right) + 10\beta_2 \lg\left(\frac{d_2}{d_0}\right) + [X_1 - X_2]_{\text{dB}}
 \end{aligned}
 \tag{4}$$

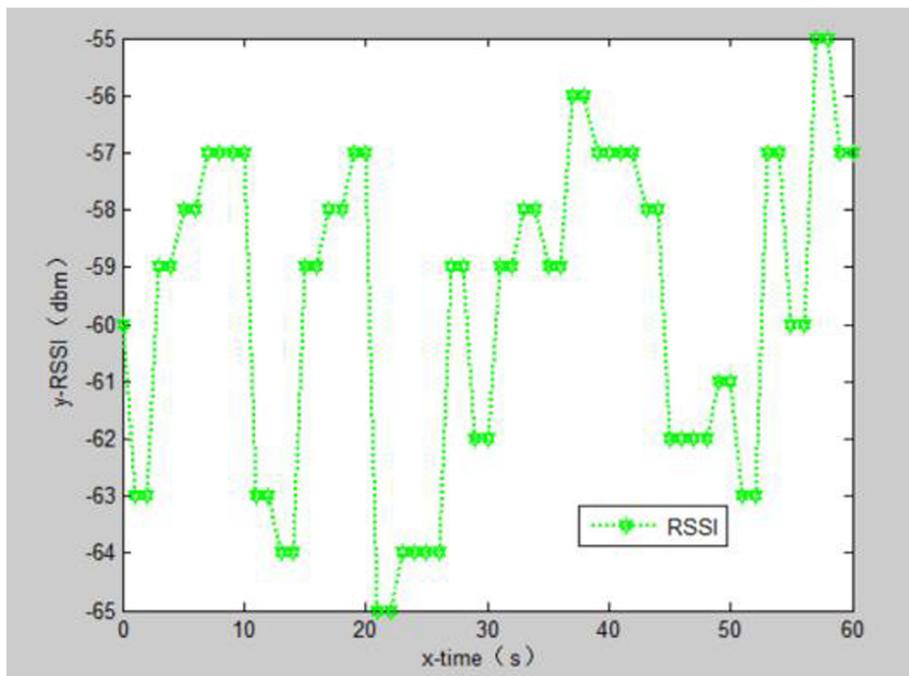


Fig. 2 RSS measured by the same device at the same location during one minute

As can be seen from formula (4), the difference of two RSS values eliminates the item of G_{MT} , that is, RSSD eliminates the impact of device heterogeneity. The difference among APs can be used as fingerprint features, for example, RSS differences, i.e., a RSSD value between two APs can be used as a fingerprint. RSSD can reflect the interval relationship between fingerprints at a specific time and location, and it is a more stable wireless signal feature than the absolute RSS value. But, for n APs, the values of independent RSSD can only be $n - 1$. This reduction in dimension may make RSSD slightly less accurate than RSS when using the same device in the offline and online stages. However, Kjargaard [45] has pointed out that the increase of RSS fingerprint dimension does not bring obvious improvement of localization accuracy when n is greater than 5. So, when using different devices, RSSD as the location fingerprint can bring significant improvement to accuracy.

WLAN-based fingerprint localization consists of two stages: the offline stage and the online stage. During the offline stage, RSS reading of each access point (AP) was collected at the uniformly distributed reference point (RP), and the RSS values were clustered and eliminated outliers. The most stable m APs and their corresponding RSS values of each cluster were extracted as the eigenvalues for the subsequent rough localization. During the online stage, first, the AP and corresponding RSS reading measured online was compared with the clustering eigenvalues in the offline stage to find the most similar

clusters and roughly locate several subsets. Then, the orthogonalization is done on the set to meet the necessary conditions for compressive sensing principle. The processed set is calculated as the RSSD between APs. Finally, the compressive sensing theory is used for accurate localization. The flow chart of the proposed scheme is shown in Fig. 3

3.1 Offline stage

In general, the localization application scenario is that the user holds a mobile device, measures RSS reading of some APs which can be sensed, and the user's location is shown on the map or illustrated as a message. The user's location estimation is done by comparing the similarity between the RSS values measured at the mobile and the fingerprint map. In this paper, the compressive sensing theory is used for reconstructing the fingerprint map. In the process of fingerprint matching, the RSSD value is used as the new fingerprint to compare the measured value to the database, to find the most matching fingerprint, and to realize the localization.

3.1.1 Fingerprint collection

In the offline phase, the original set of the RSS time samples collected from AP i in RP j are denoted as $\{\psi_i, j(\tau), \tau = 1, 2, \dots, q, q > 1\}$, where q represents the total number of time samples. The experiment result shows that the mean value of received signal strength tends to

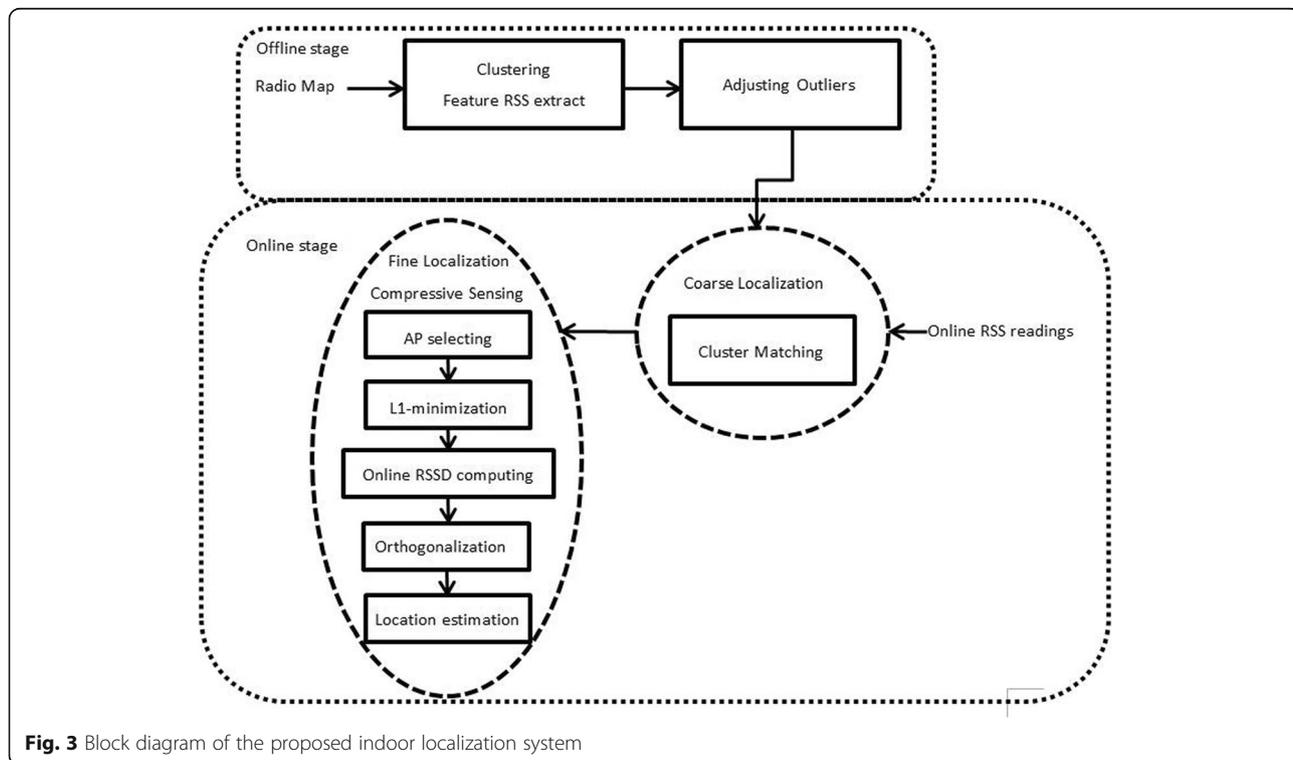


Fig. 3 Block diagram of the proposed indoor localization system

be stable when $q \geq 30$ in the experimental environment. Then, the average of the RSS time samples is computed and stored in a database, known as a radio map. Such a radio map is used for describing the RSS spatial nature of the location area and can be represented by Ψ :

$$\Psi = \begin{bmatrix} \psi_{1,1} & \psi_{1,2} & \cdots & \psi_{1,N} \\ \psi_{2,1} & \psi_{2,2} & \cdots & \psi_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{L,1} & \psi_{L,2} & \cdots & \psi_{L,N} \end{bmatrix} \quad (5)$$

where $\psi_{i,j}$ is the average of RSS readings from AP i at RP j (unit: dBm), that is, $\psi_{i,j} = 1/q \cdot \sum_{\tau=1}^q \psi_{i,j}(\tau)$, $i = 1, 2, \dots, L$, $j = 1, 2, \dots, N$, L is the total number of APs that can be detected, and N is the number of RPs. Column vector $\vec{\psi}_j$ in Ψ represents the average RSS vector that at the RP j received from L APs, is denoted as:

$$\vec{\psi}_j = [\psi_{1,j}, \psi_{2,j}, \dots, \psi_{L,j}]^T, j \in \{1, 2, \dots, N\} \quad (6)$$

where the superscript T denotes the transposition. The variable values of the RSS time sampling sequence from each AP at all RPs are stored in the database to be used in fine localization by the AP's selection mechanism subsequently. The variance vector corresponding to the RP, $j = 1, 2, \dots, N$ is defined as:

$$\vec{\Delta}_j = [\Delta_{1,j}, \Delta_{2,j}, \dots, \Delta_{L,j}]^T \quad (7)$$

where $\Delta_{i,j}$ is the unbiased estimated variance of RSS readings from AP $_i$ at RP $_j$.

$\Delta_{i,j} = 1/(q-1) \cdot \sum_{\tau=1}^q (\psi_{i,j}(\tau) - \psi_{i,j})^2$, $i = 1, 2, \dots, L$. The table composed of $\{(x_j, y_j; \vec{\psi}_j, \vec{\Delta}_j), j = 1, 2, \dots, N\}$ forms a complete fingerprint database and is stored in the server.

3.1.2 Improved fuzzy clustering method (IFCM)

Due to the indoor complex environment and complex time-varying characteristics of radio wave propagation, the online RSS is different from that in the radio map. As can be seen from Fig. 2, there is a deviation of 2–20 dBm in the RSS reading of the same AP at the same sampling point over time. To reduce the impact of time variability and remove potential outliers, the reference points in the database can be clustered. The members of the same class have the characteristics of geographic proximity and RSS fingerprint similarity. These characteristics can be used for the rough localization mechanism of the subsequent online stage, and narrow the localization area to one or more classes. Clustering can weaken the influence of time-varying and spatial interference of partial RSS, improve the efficiency of accurate localization, and reduce the time complexity of the algorithm.

There are many clustering methods, for example, the K -means clustering algorithm, FCM algorithm, ant colony algorithm, and neural network algorithm. Wu et al. [46] proposed an efficient pixel clustering-based method for mining spatial sequential patterns from serial remote sensing images. To cluster pixels into a pixel-group rapidly, the images were converted to the run-length coding schema, from which images can be overlaid with each other efficiently, to produce pixels with a sequence list. Chen et al. [47] proposed a travel time prediction system based on data clustering for waste collection vehicles. The adaptive-based clustering (ABC) methods analyze the similarities or distances between the data and cluster centers and group these data into several clusters in accordance with a threshold. He et al. [48] proposed an evolutionary K -means (EKM) algorithm, which combines K -means and genetic algorithm, solves K -means initiation problem by selecting parameters automatically through the evolution of partitions. Two aggregated consensus matrices are defined to store the clustering tendency for each pair of instances. They store the tendency that a pair of data instances should group together or group apart, respectively. The matrices were used to evaluate partitions. Liao et al. [49] proposed a tensor factorization-based user cluster (TFUC) model. The latent influence users are identified by a neural network clustering model. This model can filter the marketing users with low influence before constructing tensor, which is proven to significantly enhance the recommendation effect. Wang et al. [50] proposed two efficient co-clustering algorithms via nonnegative matrix tri-factorization. First, an effective penalty nonnegative matrix three-factorization method is introduced for high-order orthogonality constraints. Secondly, the three-factorization of symmetric penalty nonnegative matrix is used to deal with the co-aggregation problem in the extraction of the sample similarity matrix.

FCM algorithm is the commonly used fuzzy clustering method, but this algorithm is sensitive to the initial value and easy to fall into local extreme value, it is difficult to get the global optimal solution. Unlike the traditional FCM algorithm, the IFCM clustering algorithm does not need to determine the clustering center in advance. It takes all n sample points as the candidate clustering center and gives each sample point a real value, which is called the biased parameter. The larger parameters are more likely to be used as clustering centers. At the same time, the selection of biased parameters will also affect the number of clustering. The algorithm establishes similarity information between each sample point and other sample points, maximizes the fitness function in the sample competition through the message iteration of the loop between samples, and finally forms the class center and class members. Because the algorithm does

not need to randomly select the class center and the convergence speed is fast, it is applied to the data pre-processing in the offline phase of the localization system.

$\text{Sim}(i, j)$ represents the similarity between the reference point RP_j and the reference point RP_i . The received signal strength vector at the reference point RP_j can be approximately expressed as $\vec{\psi}_j + \varepsilon_j$, where ε_j is the measurement noise, approximate Gaussian distribution under certain conditions. Therefore, Euclidean distance can be used as a judgment basis to measure the similarity between RSS vectors of reference points. The similarity between reference points is defined as:

$$\text{Sim}(i, j) = -\left\| \vec{\psi}_i - \vec{\psi}_j \right\|^2, \forall i, j \in \{1, 2, \dots, N\} \quad (8)$$

where $\text{Sim}(k, k)$, $k = 1, 2, \dots, N$ is defined as the self-similarity function, and the value is assigned as the bias parameter to indicate the possibility of the reference point RP_k becoming the class center. Since there is no prior knowledge, all reference points are likely to become class centers. The bias parameters of each reference point are defined as a function of the median similarity value, which can be expressed as:

$$b = \alpha \cdot \text{median}\{\text{Sim}(i, j), \forall i, j \in \{1, 2, \dots, N\}, j \neq i\} \quad (9)$$

where α is the empirical value obtained during the experiment process, which is used for selecting a moderate number of classes.

The core of the improved FCM clustering algorithm is the transmission of messages between two reference points:

Each RP is assigned the same offset parameter to form the affiliation matrix $a[N][N]$, whose initial value is a function of the median value of $\text{Sim}(i, j)$.

Attractiveness messages are the adaptability of one reference point to be the class center of other reference points, and affiliation messages are the affiliation degree of a reference point, which becomes a class member centered on another reference point. IFCM clustering is to continuously search for evidence from data and carry out message transmission in a cyclic and iterative way, to produce a high-quality class center and assign a classification center to every member.

3.1.2.1 Attractiveness message $a(i, j)$ Attractiveness message $a(i, j)$ is sent from the reference point RP_i to RP_j , reflecting the cumulative evidence which RP_j as a class attracts to RP_i under the influence of considering another candidate reference point besides RP_j , is denoted as:

$$a(i, j) = \text{Sim}(i, j) - \max\{t(i, j') + \text{Sim}(i, j'), j' \neq j\}, \quad (10)$$

where $t(i, j')$ are affiliation messages.

3.1.2.2 Affiliation message $t(i, j)$ Affiliation message $t(i, j)$ is sent from candidate class center RP_j to RP_i , reflecting the accumulated evidence that under the influence of other class members taking RP_j as the center, RP_i believes that RP_j is belonging to its class center, they are denoted as:

$$t(i, j) = \min\left\{0, a(j, j) + \sum_{i' \neq i, j} \max\{0, a(i', j)\}\right\}, i' \neq i \quad (11)$$

The self-affiliation message $t(j, j)$ reflects the accumulated evidence of RP_j as the class center after the positive attraction values are sent from other reference points besides RP_j to RP_j . It is denoted as:

$$t(j, j) = \sum_{i' \neq j} \max\{0, a(i', j)\} \quad (12)$$

Class centers compete through passing two messages between the pairs of reference points. For every RP_i , calculating $j' = \text{argmax}_{j \in \{1, 2, \dots, N\}} \{t(i, j) + a(i, j)\}$, if $j' = i$, then RP_i is selected as the class center. On the contrary, RP_j is going to be the class center of RP_i . The messages are passed recursively between pairs of RPs within each radio map, and the above updating rules are followed until the appropriate number of class convergences, the corresponding class center is formed.

In general, the reference points located at the center of the cluster usually have a bigger sum of attraction to other reference points. Therefore, they are more likely to become the class center. On the contrary, for those located at the border of the cluster, the possibility of becoming the class center is low due to smaller sums of attraction.

3.2 Online stage

Define the RSS value measured by the user's handheld terminal device at any unknown location as:

$$\vec{\phi}_r = [\phi_{1,r}, \dots, \phi_{L,r}]^T \quad (13)$$

where $\phi_{\kappa, r}$ is the online average RSS. The online localization stage is divided into a coarse localization stage and fine localization stage. The coarse localization first compares the similarity between the RSS vectors $\vec{\phi}_r$ and every class center in the offline database, and an appropriate threshold is set to narrow the localization area with a sub-set of the reference point, which not only reduces the computational complexity of the localization algorithm but also can remove the errors caused by outliers at a distance. Then, select the most powerful m APs from the RSS vector $\vec{\phi}_r$, measured online, conducting the AP selection of the data in the offline database and select the same m APs to participate in the calculation. To eliminate device heterogeneity, the subtraction between the first AP and the other $m - 1$

APs was performed to generate a new RSSD matrix for localization.

3.2.1 Coarse localization

According to the definition of similarity function, the set of matching class center is denoted as:

$$\text{Sim}_{\text{mt}} = \{\text{Sim}(r, j) > \beta, j \in H\} \quad (14)$$

where H is the number of clusters, β is the threshold for controlling the number of matching classes, which is defined as the function of maximum similarity and minimum similarity, and is denoted as:

$$\beta = \max\{\text{Sim}(r, j)\} + \beta_1 \min\{\text{sim}(r, j)\} \quad (15)$$

where β_1 is usually a real number between 0.3 and 0.5.

After the matching class is determined, the RSS vector set received by the members of the matching class is denoted as a $L \times \tilde{N}$ matrix $\tilde{\Psi}$:

$$\tilde{\Psi} = [\tilde{\psi}_{i,j}, i = 1, 2, \dots, L, \forall j \in C] \quad (16)$$

which means $\tilde{\Psi} = \begin{bmatrix} \tilde{\psi}_{1,1} & \tilde{\psi}_{1,2} & \dots & \tilde{\psi}_{1,\tilde{N}} \\ \tilde{\psi}_{2,1} & \tilde{\psi}_{2,2} & \dots & \tilde{\psi}_{2,\tilde{N}} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\psi}_{L,1} & \tilde{\psi}_{L,2} & \dots & \tilde{\psi}_{L,\tilde{N}} \end{bmatrix}$, where

\tilde{N} is the number of matching class members, and means $\tilde{N} = |C|$.

3.2.1.1 AP selection mechanism for the strongest RSS

According to the measured value of online RSS, m APs with the strongest RSS are selected for calculation. Sort the online RSS received from L APs in descending order, select the first M APs, and complete the matrix c line by line. Since the user receives different signals from different APs in different positions, the matrix $\tilde{\Psi}$ needs to be filled in dynamically.

$$\tilde{\Psi} = \begin{bmatrix} \tilde{\psi}_{1,1} & \tilde{\psi}_{1,2} & \dots & \tilde{\psi}_{1,N} \\ \tilde{\psi}_{2,1} & \tilde{\psi}_{2,2} & \dots & \tilde{\psi}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\psi}_{M,1} & \tilde{\psi}_{M,2} & \dots & \tilde{\psi}_{M,N} \end{bmatrix} \quad (17)$$

3.2.1.2 Received signal strength difference (RSSD) conversion

After determining the matrix $\tilde{\Psi}$, the subtraction between the other $m - 1$ APs and the first AP is calculated to form the matrix $\tilde{\Psi}'$.

$$\tilde{\Psi}' = \begin{bmatrix} \tilde{\psi}'_{1,1} & \tilde{\psi}'_{1,2} & \dots & \tilde{\psi}'_{1,N} \\ \tilde{\psi}'_{2,1} & \tilde{\psi}'_{2,2} & \dots & \tilde{\psi}'_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\psi}'_{M-1,1} & \tilde{\psi}'_{M-1,2} & \dots & \tilde{\psi}'_{M-1,N} \end{bmatrix} \quad (18)$$

where $\tilde{\psi}'_{i,j}$ is the difference between the signal strength received from AP_i and AP_1 at the reference point RP_j , and is denoted as:

$$\tilde{\psi}'_{i,j} = \tilde{\psi}_{i,j} - \tilde{\psi}_{1,j} \quad (19)$$

Let M represent the total number of the strongest AP selected in the localization area and \tilde{N} represent the total number of reference points in the matching class.

The RSSD vector received online is defined as $\vec{\phi}'_r$,

$$\vec{\phi}'_r = [\phi'_{1,r}, \dots, \phi'_{M-1,r}]^T \quad (20)$$

where $\phi'_{i,r}$ represents the RSSD between AP_i and AP_1 at the request location, that is:

$$\phi'_{i,r} = \phi_{i,r} - \phi_{1,r} \quad (21)$$

3.2.2 Fine localization

3.2.2.1 Sparse characteristics and problem modeling

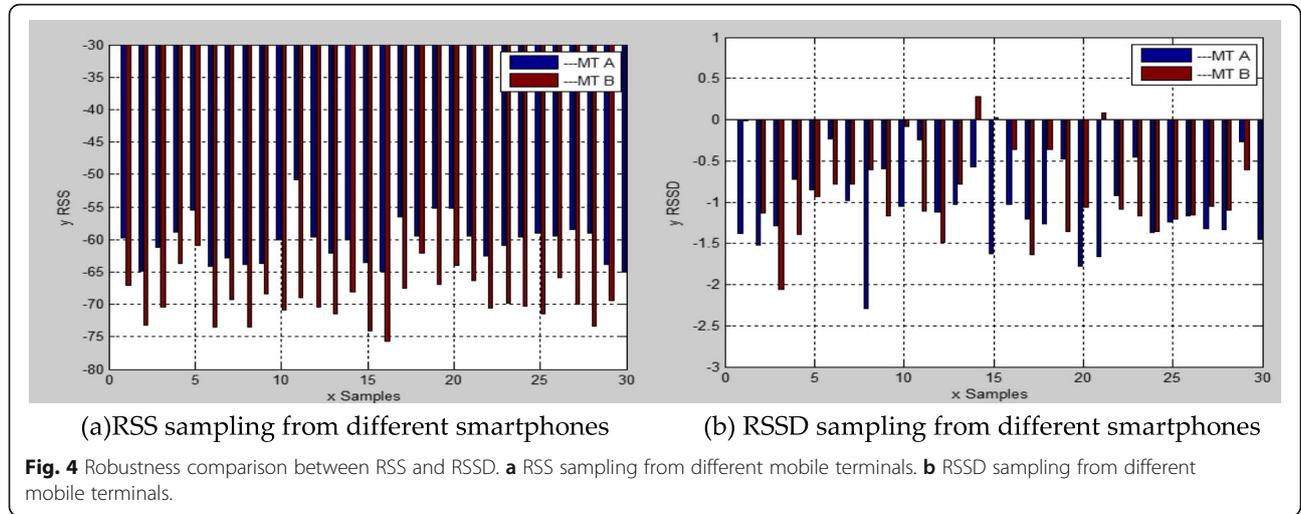
To apply compressive sensing, sparse characteristics and irrelevancy are required. Firstly, for the localization system, the spatial position of the user at a specific time is unique and sparse. Ideally, suppose that the user just stands on a RP, so θ can be defined as a vector of $\tilde{N} \times 1$, and each element represents a reference point in the spatial domain selected in rough localization. Thus, the user's position can be accurately represented by a 1-dimensional sparse vector, that is, the corresponding reference point position that the user is standing on is 1, and the other positions are 0:

$$\theta = [0, \dots, 0, 1, 0, \dots, 0]^T \quad (22)$$

Set y as the measurement vector of online RSS, y' as the vector composed of the RSSD between AP_i and AP_1 , and y' can be further described as:

$$y' = \tilde{\Psi}' \theta + \delta \quad (23)$$

where $\tilde{\Psi}' = \Phi \tilde{\Psi}$; the definition of $\tilde{\Psi}'$ is shown in formulas (18) and (19). Φ is $M \times L$ matrix and is a selection on the measurement vector y . Each row of Φ is a $1 \times L$ vector, only $\phi(r) = 1$, where r is the index of selected AP. $\tilde{\Psi}$ is a $L \times \tilde{N}$ matrix, is a subset of the offline fingerprint library defined in (16), each column represents the RSS



vector of a selected class member in the coarse localization region. δ is unknown environmental noise. In the ideal environment without considering noise, $\vec{\phi}_r'$ is a column about $\tilde{\Psi}'$ corresponding to an index position of θ . In this case, the relationship between the measurement vector y and the online received signal strength vector ϕ_r is:

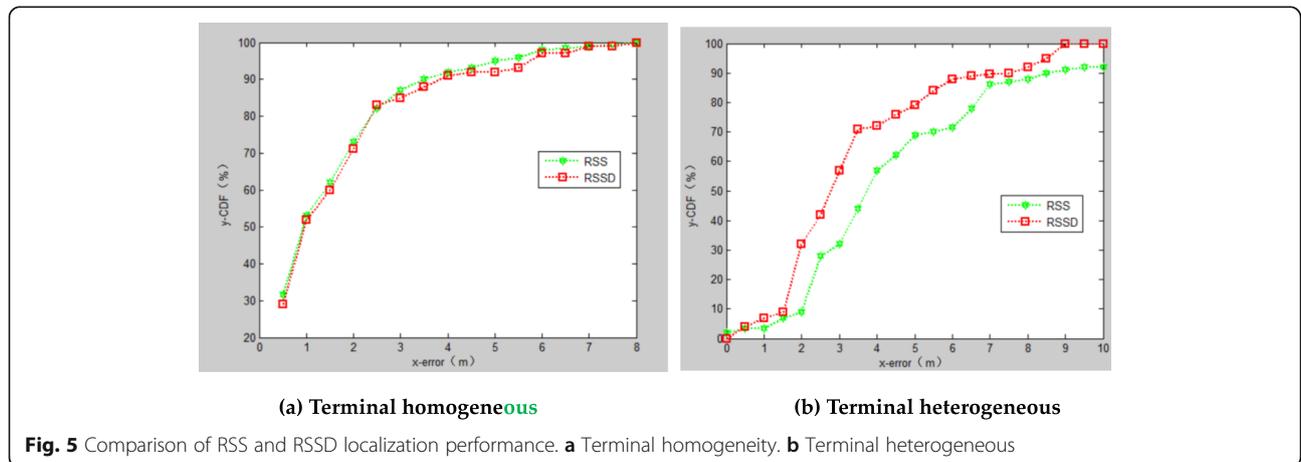
$$y = \Phi\phi_r \tag{24}$$

The purpose of localization is to reconstruct the 1-dimensional sparse vector θ from formula (23) according to the measurement vector difference y' at the moment t , y' is calculated by formulas (21) and (20) after the AP selection mechanism.

3.2.2.2 Orthogonalization In addition to the sparse characteristics, the uncorrelated or weak correlation matrix $\tilde{\Psi}'$ is an important feature to ensure that the

linear projection of signals can maintain the original structure of signals, which become another important guarantee for sparse signals reconstruction using the compressive sensing theory. The restricted isometry property (RIP) condition of the matrix $\tilde{\Psi}'$ is given and proved by Candes and Tao [51], and it is a necessary condition using the compressive sensing theory to realize sparse signals reconstruction through ℓ_1 -minimization [52, 53] approximating ℓ_0 -minimization. Therefore, orthogonalizing the matrix $\tilde{\Psi}'$ is necessary.

Orthogonalize y defined in formula (24), that is, $z = Ty$. Define $T = QR^+$. $R = \tilde{\Psi}'$, $Q = orth(R^T)^T$, represents the orthogonalization of the matrix R , R^+ is the pseudo-inverse matrix of R . Therefore, the localization problem can be described as the minimization problem of the following ℓ_1 -norm:



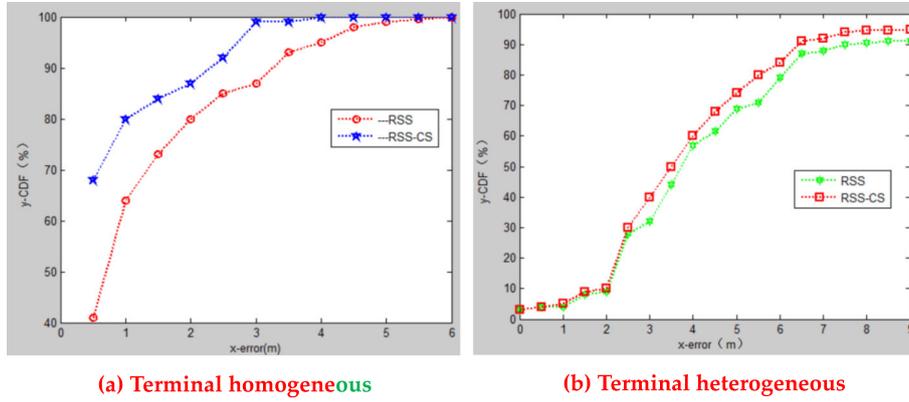


Fig. 6 Comparison of RSS-based using compressive sensing localization performance. **a** Terminal homogeneity. **b** Terminal heterogeneous

$$\hat{\theta} = \arg \min_{\theta \in R^N} \|\theta\|_1, s.t. z = Q\theta + \delta' \quad (25)$$

$z = Ty = QR^+ \cdot (R\theta + \delta) = QR^+ \cdot R\theta + \delta' = Q\theta + \delta'$ $\tilde{\Psi}'$ is known, Q meets the RIP characteristic [54]. Since θ has a sparse nature, according to the theory of compressive sensing theory [53, 55–58], if the number of APs M meets $M = O(\log \tilde{N})$, the location indicator θ can be well recovered from z with very high probability by solving the following ℓ_1 -minimization problem:

When the user happens to be located at a reference point, it is ideally reconstructed into a 1-dimensional sparse vector, the value of the index position is 1, while the other positions are 0. However, in practice, the user does not necessarily happen to be at a certain reference point, so the reconstruction of $\hat{\theta}$ is not necessarily a 1-dimensional sparse vector. Some larger values are taken at the index positions of a few candidate reference points, and the remaining index positions are approximately 0. To some extent, these non-zero values between (0, 1) reflect the possibility of corresponding candidate reference points as position estimation. Therefore, the threshold σ is set, and the reference point corresponding to the coefficient greater than the threshold σ is taken as the second-level candidate reference point and is denoted as Y ,

$$Y = \{n | \hat{\theta}(n) > \sigma\} \quad (26)$$

Therefore, the linear combination of these second-order reference points with weights will be used as the final position estimation for accurate localization.

$$(\hat{x}, \hat{y}) = \sum_{n \in Y} \mu_n \cdot (\hat{x}_n, \hat{y}_n) \quad (27)$$

where $\mu_n = \hat{\theta}(n) / \sum_{i \in Y} \hat{\theta}(i)$. The localization is completed.

4 Experimental results and discussion

4.1 RSSD robustness analysis

Thirty locations were randomly selected and sampled with two different mobile terminals for signal strength. RSS and RSSD values of each location were compared, as shown in Fig. 4. Figure 4a is the RSS sequence of the same AP at 30 sampling points of two different mobile terminals. Figure 4b is the RSSD sequence of the difference in signal strength for a pair of APs in two different terminals at 30 sampling points.

It can be seen that at each sampling point, the RSS values obtained by different mobile terminals are different, with a maximum difference of 14.31 dBm, while the RSSD values obtained by different mobile terminals maintain a relative consistency. Although there is a difference between the two RSSD values at the same sampling point, the maximum difference is only 1.74 dBm. That RSSD had better robustness compared with RSS is consistent with the previous analysis.

4.2 Comparison of fingerprint location accuracy of RSS and RSSD

In practice, terminals used in localization are mostly different from that used in fingerprint database construction. At the same time, the number of AP used for localization is also a major factor affecting the localization accuracy. Therefore, in the process of the experiment, two cases of terminal homogeneity and heterogeneity were tested, respectively. Using the traditional KNN algorithm, when the same type of terminals is used for the offline fingerprint database construction and the online localization, the localization accuracy of RSSD is slightly lower than that of the RSS when 6 APs are selected. When 10 APs are selected, their performance is similar, and both of them can achieve the localization accuracy within 3 m in case of over 87%. However, in the case of different types of terminals, the performance of the RSS decreases significantly, and the localization accuracy within 3 m is only 32%, while

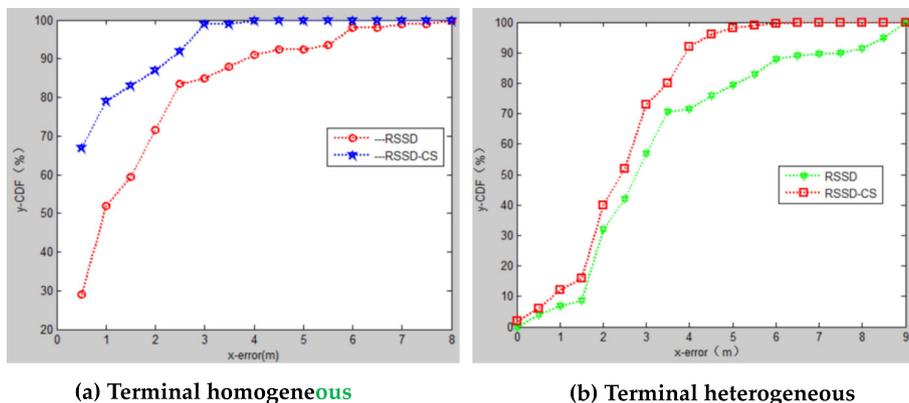


Fig. 7 Comparison of RSSD-CS localization performance. **a** Terminal homogeneous. **b** Terminal heterogeneous

the localization accuracy within 3 m of RSSD reaches 58%, 85% of the probability can reach the localization accuracy within 5 m, and the localization accuracy can be improved by more than 20% as shown in Fig. 5.

It can be seen that the localization accuracy in the case of terminal heterogeneity is significantly lower than that in the case of homogeneity. RSS-based localization decreased by more than 50%, RSSD-based decreased by nearly 30%. But in the practical application, the situation of terminal heterogeneity is universal, so it is urgent to improve the performance of the localization system.

4.3 Using compressive sensing localization performance

In the case of terminal homogeneous, the RSS-based using compress sensing localization accuracy can reach 100% within 3 m, 80% within 1 m and 67% within 0.4 m, with the performance improved by nearly 20%. In the case of terminal heterogeneous, the localization accuracy within 3 m can be achieved at 40% and within 4 m at 60%, with the performance improved by nearly 3.8%. It can be seen that compressive sensing has no obvious effect on device heterogeneity, which is shown in Fig. 6.

In the case of terminal homogeneous, the localization accuracy of the RSSD-based using compress sensing can reach 100% within 3 m, 79% within 1 m and 67% within 0.4 m, with the performance improved by nearly 20%. In the case of terminal heterogeneous, the localization accuracy within 3 m can be achieved at 73% and within 4 m at 92%. It can be seen that RSSD-CS fusion algorithm

greatly improves the positioning accuracy in the case of homogeneous and heterogeneous terminals, which is shown in Fig. 7.

The performance of RSSD-CS algorithm is verified through experiments, and the specific summary is shown in Tables 1 and 2.

5 Conclusion

This paper proposes an RSSD-CS indoor localization algorithm to address the terminal heterogeneity, which could improve the indoor localization accuracy. Firstly, an IFCM algorithm is used for clustering map data to eliminate outliers in the offline phases, narrow the search range for online matching, and reduce time complexity. Then, during the online phases, RSSD values of APs sensed and selected at location points and matching matrices on the map are calculated orthogonally to meet the constraints of CS. Finally, CS theory is used for estimating the user’s location. The experimental results demonstrate that the proposed localization method leads to substantial improvements in localization accuracy over the widely used traditional fingerprinting methods. In particular, the positioning accuracy of this method is improved by 20.5% and 15.6% compared to SSD and received signal strength and compressive sensing (RSS-CS) algorithm.

Future research of this work may have three directions. First, the proposed IFCM algorithm has a higher time complexity, which requires a longer time to calculate when the number of samples is large. While this

Table 1 Different percentile values concerning the positioning error under terminal homogeneous

Percentile values	RSSD-CS positioning	RSS positioning	RSSD positioning	RSS-CS positioning
50%	< 0.25 m	< 1.0 m	< 1.0 m	< 0.35 m
70%	< 0.6 m	< 1.9 m	< 1.9 m	< 0.55 m
90%	< 2.3 m	< 3.5 m	< 3.4 m	< 2.3 m

Table 2 Different percentile values concerning the positioning error under terminal heterogeneous

Percentile values	RSSD-CS positioning	RSS positioning	RSSD positioning	RSS-CS positioning
50%	< 2.4 m	< 3.7 m	< 2.8 m	< 3.5 m
70%	< 2.9 m	< 5.9 m	< 3.5 m	< 4.7 m
90%	< 3.9 m	< 8.9 m	< 7.5 m	< 6.3 m

work only improves the shortcomings of the clustering algorithm in terms of its sensitivity to initial values, an in-depth research for other clustering algorithms could be conducted in the future. Second, the calculation of RSSD values that is conducted in online stages increases the computational complexity. If RSSD values of map data can be obtained offline in the future, the timeliness of location could be guaranteed. Third, the localization accuracy is also related to the network element layout. If the layout of network elements can be optimized in the future, which could ensure that 5–7 first-path signals can be received at any point in the positioning space, the effect of multi-path can be effectively avoided and the positioning accuracy can be further improved.

Abbreviations

WLAN: Wireless local area network; RSSI: Received signal strength indication; RSS: Received signal strength; RSSD: Received signal strength difference; CS: Compressive sensing; RSSD-CS: Received signal strength difference and compressive sensing; SSD: Signal strength difference; RSS-CS: Received signal strength and compressive sensing; Wi-Fi: Wireless fidelity; LBS: Location-based service; RP: Reference point; AP: Access point; RF: Radio frequency; FCM: Fuzzy clustering method; IFCM: Improving fuzzy clustering method; ABC: Adaptive-based clustering; EKM: Evolutionary *K*-means; TFUC: Tensor factorization-based user cluster; KNN: *K*-nearest neighbor; CF: Curve fitting; MT: Mobile terminal

Authors' contributions

YXM contributed to the investigation, methodology, draft manuscript writing, manuscript reviewing, editing, software and hardware development, simulations, result in analysis, and reviewing. WHQ and WJQ contributed to the overall design and network element optimization layout model. The authors read and approved the final manuscript.

Funding

This research was supported by National Natural Science Foundation of China (no. 61872104 and 61901134), the National Science and Technology Major Project of China (no. 2016ZX03001023-005), the China Postdoctoral Science Foundation (no. 2019M651264), Natural science foundation of Heilongjiang Province of China (no. 2015023) and the Basic Business Project in Education Department of Heilongjiang Province of China (no. 135209244).

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on a reasonable request.

Competing interest

The authors declare that they have no competing interests.

Author details

¹College of Computer Science and Technology, Harbin Engineering University, Harbin, People's Republic of China. ²College of Computer and Control Engineering, Qiqihar University, Qiqihar, People's Republic of China. ³College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin, People's Republic of China.

Received: 31 July 2019 Accepted: 6 March 2020

Published online: 03 April 2020

References

1. Castro P, Chiu P, Kremenek T, et al. A probabilistic room location service for wireless networked environments [J]. 2001.
2. C. Wu, Z. Yang, Y. Liu, et al., WILL: wireless indoor localization without site survey [J]. *IEEE Transactions on Parallel & Distributed Systems* **24**(4), 839–848 (2013)
3. M. Bshara, U. Orguner, F. Gustafsson, et al., Fingerprinting localization in wireless networks based on received-signal-strength measurements: a case study on WiMAX networks [J]. *IEEE Transactions on Vehicular Technology* **59**(1), 283–294 (2010)
4. Shih CY, Chen LH, Chen GH, et al. Intelligent radio map management for future WLAN indoor location fingerprinting [C] // *Wireless Communications & Networking Conference*. IEEE, 2012.
5. S.H. Fang, T.N. Lin, P.C. Lin, Location fingerprinting in a decorrelated space [J]. *IEEE Transactions on Knowledge and Data Engineering* **20**(5), 685–691 (2008)
6. L. Wei, Research on the location of WIFI location and the design of simulators. Master's thesis of southwest jiaotong university, 5–14 (2012)
7. Li B, Wang Y, Lee H K, et al. Method for yielding a database of location fingerprints in WLAN [J]. *IEE Proceedings-Communications*, 2005, 152(5):580-0.
8. B. Wang, S. Zhou, W. Liu, Y. Mo, Indoor localization based on curve fitting and location search using received signal strength. *IEEE transactions on industrial electronics*. **62**(1), 572–582 (2015)
9. Dai-Jun X, Fan-Zeng K, Han-Ying Hu. Research on robustness of location fingerprint under terminal heterogeneity [J]. *Computer Engineering*, 2014.
10. Sharma P, Chakraborty D, Banerjee N, et al. KARMA: Improving WiFi-based indoor localization with dynamic causality calibration [C] // *SECON 2014*. IEEE, 2014.
11. K. Kaemarungsi, *Efficient design of indoor positioning systems based on location fingerprinting* (Wireless Networks, Communication and Mobile Computing, 2004), pp. 18–24
12. K. Pahlavan, X. Li, J.P. Makela, Indoor geolocation science and technology [J]. *IEEE Communications Magazine* **40**(2), 112–118 (2002)
13. Seshadri V, Zaruba G V, Huber M. A Bayesian sampling approach to in-door localization of wireless devices using received signal strength indication [C] // *Pervasive Computing and Communications, 2005. PerCom 2005. Third IEEE International Conference on*. IEEE, 2005, 75–84.
14. P. Bahl, V.N. Padmanabhan, Radar: An in-building rf-based user location and tracking system. *INFOCOM*, 775–784 (2000)
15. M. Brunato, R. Battiti, Statistical learning theory for location fingerprinting in wireless LANs [J]. *Computer Networks* **47**(6), 825–845 (2005)
16. R. Battiti, V.A. Le NT, *Location-aware computing: a neural network model for determining location in wireless lans* (tech. rep., Technical Report DIT-02-0083, Department of Information and Communication Technology, University of Trento, Italy, 2002)
17. C. Nerguizian, C. Despins, S. Affès, Geolocation in mines with an impulse response fingerprinting technique and neural networks [J]. *IEEE Transactions on Wireless Communications* **5**(3), 603–611 (2006)
18. M.A. Youssef, A. Agrawala, A.U. Shankar, in *Pervasive Computing and Communications*. Wlan location determination via clustering and probability distributions (2003), pp. 143–150
19. D. Fox, J. Hightower, L. Liao, et al., Bayesian filtering for location estimation [J]. *IEEE Pervasive Computing* **2**(3), 24–33 (2003)
20. Z.L. Wu, C.H. Li, J. Ng, et al., Location estimation via support vector regression [J]. *IEEE Transactions on Mobile Computing* **6**(3), 311–321 (2007)
21. A. Kushki, K.N. Plataniotis, A.N. Venetsanopoulos, Kernel-based positioning in wireless local area networks [J]. *IEEE Transactions on Mobile Computing* **6**(6), 689–705 (2007)

22. J.F. Pan, J. Kwok, Q. Yang, et al., Accurate and low-cost location estimation using kernels. *International Joint Conference on Artificial Intelligent*, 1366–1370 (2005)
23. J.J. Pan, J.T. Kwok, Q. Yang, et al., Multidimensional vector regression for accurate and low-cost location estimation in pervasive computing [J]. *IEEE Transactions on Knowledge and Data Engineering* **18**(9), 1181–1193 (2006)
24. Y. Cheng, H. Jiang, F. Wang, et al., Using high-bandwidth networks efficiently for fast graph computation [J]. *IEEE Transactions on Parallel and Distributed Systems*, 1–1 (2018)
25. S.H. Fang, C.H. Wang, Y. Tsao, Compensating for orientation mismatch in robust Wi-Fi localization using histogram equalization. *IEEE Transactions on Vehicular Technology*. **64**(11), 5210–5220 (2015)
26. Alejandro A. Antennas and Propagation for Wireless Communication Systems: 2nd Edition[M]. (1999)
27. S.H. Fang, C.H. Wang, A novel fused positioning feature for handling heterogeneous hardware problem [J]. *IEEE Transactions on Communications* **63**(7), 2713–2723 (2015)
28. C.H. Chen, An arrival time prediction method for bus system [J]. *IEEE Internet of Things Journal*, 1–1 (2018)
29. S.H. Fang, C.C. Chuang, C. Wang, Attack-resistant wireless localization using an inclusive disjunction model [J]. *IEEE Transactions on Communications* **60**(5), 1209–1214 (2012)
30. W.J. Chang, J.H. Tarng, Effects of bandwidth on observable multipath clustering in outdoor/indoor environments for broadband and ultra-wide-band wireless Systems [J]. *IEEE Transactions on Vehicular Technology* **56**(4), 1913–1923 (2007)
31. H. Hashemi, Impulse response modeling of indoor radio propagation channels [J]. *IEEE J Selected Areas in Commun* **11**(2), 594–606 (1993)
32. A.K.M. Mahtab Hossain, Y. Jin, W.S. Soh, et al., SSD: a robust RF location fingerprint addressing mobile devices' heterogeneity [J]. *IEEE Transactions on Mobile Computing* **12**(1), 65–77 (2013)
33. N. Brouwers, M. Zuniga, K. Langendoen, Incremental Wi-Fi scanning for energy-efficient localization[C]//2014 IEEE International Conference on Pervasive Computing and Communications (PerCom). IEEE, (2014) <https://doi.org/10.1109/PerCom.2014.6813956>
34. A. Haeberlen, Practical robust localization over large-scale 802.11 wireless networks[C]// International Conference on Mobile Computing and Networking. ACM, 70–84 (2004)
35. M.B. Kjaergaard, Indoor location fingerprinting with heterogeneous clients [J]. *Pervasive and Mobile Computing* **7**(1), 31–43 (2011)
36. C. Laoudias, D. Zeinalipour-Yazti, C. Panayiotou, et al., Crowdsourced indoor localization for diverse devices through radiomap fusion [J], in *Proc. IPIN* **25**, 28–34 (2013)
37. L. Li, G. Shen, C. Zhao, et al., Experiencing and handling the diversity in data density and environmental locality in an indoor positioning service[C]// International Conference on Mobile Computing & Networking ACM, 459–470 (2014)
38. C. Feng, W.S.A. Au, S. Valaee, et al., Received-signal-strength-based indoor positioning using compressive sensing [J]. *IEEE Transactions on Mobile Computing* **11**(12), 1983–1993 (2012)
39. W. Xiaoyang, T. Wenyuan, O. Chung-Ming, et al., On the dynamic RSS feedbacks of indoor fingerprinting databases for localization reliability improvement [J]. *Sensors* **16**(8), 1278–1294 (2016)
40. H. Zheng, M. Gao, Z. Chen, et al., An adaptive sampling scheme via approximate volume sampling for fingerprint-based indoor localization [J]. *IEEE Internet of Things Journal*, 1–1 (2019)
41. L. Wu, C.H. Chen, Q.S. Zhang, A mobile positioning method based on deep learning techniques. *Electronics* **8**(1), 59 (2019)
42. R. Cheng, Y. Song, D. Chen, et al., Intelligent positioning approach for high speed trains based on ant colony optimization and machine learning algorithms [J]. *IEEE Transactions on Intelligent Transportation Systems*, 1–10 (2018)
43. G.Y. Chen, M. Gan, C.L.P. Chen, et al., A two-stage estimation algorithm based on variable projection method for GPS positioning [J]. *IEEE Transactions on Instrumentation and Measurement* **67**(11), 1–8 (2018)
44. X. Dai-Jun, K. Fan-Zeng, H. Han-Ying, Research on robustness of location fingerprint under terminal heterogeneity [J]. *Computer Engineering* **5**(40), 81–85 (2014)
45. M.B. Kjaergaard, Hyperbolic location fingerprinting: a calibration-free solution for handling differences in signal strength[C]// Proc. Sixth Annual IEEE International Conference on Pervasive Computing and Communications. IEEE, 17–21 (2008) <https://doi.org/10.1109/PERCOM.2008.75>
46. X. Wu, X. Zhang, An efficient pixel clustering-based method for mining spatial sequential patterns from serial remote sensing images [J]. *Computers & Geosciences* **124**, 128–139 (2019)
47. Ch.H. Chen, F.J. Hwang, H.Y. Kung, Travel Time Prediction System Based on Data Clustering for Waste Collection Vehicles, *IEICE TRANS. INF. & SYST.* **E102–D**(7), 1374–1383 (2019)
48. H. Zhenfeng, Y. Chunyan, Clustering stability-based evolutionary K-means [J]. *Soft Computing* (2018)
49. Liao X W , Zhang LY, Wei J J , et al. Recommending mobile microblog users via a tensor factorization based on user cluster approach, *Hindawi Wireless Communications and Mobile Computing Volume 2018*, Article ID 9434239, 11 pages <https://doi.org/10.1155/2018/9434239>
50. S. Wang, A. Huang, Penalized nonnegative matrix tri-factorization for co-clustering [J]. *Expert Systems with Applications* **78**, 64–73
51. E. Candes, J. Romberg, Sparsity and incoherence in compressive sampling [J]. *Inverse Problems* **23**(3), 969–985 (2007)
52. S.S. Chen, D.M.A. Saunders, Atomic decomposition by basis pursuit [J]. *SIAM Review* **43**(1), 129–159 (2001)
53. E.J. Candes, M.B. Wakin, S.P. Boyd, Enhancing sparsity by reweighted ℓ_1 minimization [J]. *Journal of Fourier Analysis & Applications* **14**(5–6), 877–905 (2008)
54. R. Baraniuk, M. Davenport, R. Devore, et al., A simple proof of the restricted isometry property for random matrices [J]. *Constructive Approximation* **28**(3), 253–263 (2008)
55. Zhang Y. Theory of compressive sensing via ℓ - minimization: a non-rip analysis and extensions, Technical Report TR08-11, Rice CAAM Dept., 2008.
56. E.J. Candes, J.K. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements [J]. *Communications on Pure & Applied Mathematics* **59**(8), 1207–1223 (2006)
57. J. Wu, X. Ma, Y. Yin, et al., A novel algorithm to mitigate the effect of clipping in orthogonal frequency division multiplexing underwater communication acoustic sensor system [J]. *International Journal of Distributed Sensor Networks* (2017)
58. C.H. Chen, F.Y. Song, et al., in *Physical A: Statistical Mechanics and its Applications*. A probability density function generator based on neural networks, vol 541 (2020)

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)