RESEARCH

Open Access

Cost-efficient BLE fingerprint database construction approach via multi-quadric RBF interpolation

Liangbo Xie[1,2], Xiaoxiao Jin[1,2]*, Mu Zhou[1,2], Yue Wang[1,2] and Zengshan Tian[1,2]

Abstract

The demand for indoor localization is becoming urgent, but the traditional location fingerprint approach takes a lot of manpower and time to construct a fine-grained location fingerprint database. To address this problem, we propose to use the approach of combining dynamic collection of fingerprint samples with Radial Basis Function (RBF) interpolation. Specifically, the raw sparse fingerprint database is constructed from a small number of fingerprints collected on a few paths, in which the pedestrian track correction algorithm improves the validity and accuracy of the sparse fingerprint database. Then, the RBF interpolation approach is applied to enrich the sparse fingerprint database, in which the Genetic Algorithm (GA) is used to optimize the free shape parameter and the cut-off radius is determined according to the experimental results. Extensive experiments show that the proposed approach guarantees high interpolation and localization accuracy and also significantly reduces the effort of manual collection of fingerprint samples.

Keywords: Bluetooth Low Energy, Indoor localization, Location fingerprint database, Radial basis function, MEMS

1 Introduction

With the rapid development of intelligent terminals and communication technologies, location-based services have become an indispensable and important element in daily life. Innovative services such as Didi taxi and Mobike shared bicycles are closely related to the development and application of localization technology. And intelligent transportation is based on the Internet of Things (IoT) [1], 5G [2], and other technologies, which provide enormous connections of devices and sensors with applications. Satellite communications keep step with the quick development of wireless terrestrial communications [3, 4], such as Global Positioning System (GPS) [5], and can provide good localization services. However, due to the blockage of buildings in the indoor environment, it is difficult to receive satellite signals. Therefore, the upsurge of research and positioning has been transferred to the indoor localization. The existing indoor localization systems are mainly based on Wireless Local Area Network (WLAN) [6], Bluetooth Low Energy (BLE) [7], Micro Electro Mechanical Systems (MEMS) [8], Radio Frequency Identification (RFID) [9], and Ultra Wideband (UWB) [10]. Working in the same frequency band as WLAN, Bluetooth technology has attracted extensive attention of researchers. With the increasing integration degree of Bluetooth modules, a variety of Bluetooth-based indoor localization methods [11] emergy, such as range detection-based localization method, signal strength-based localization method, and Cell-ID-based localization method. The introduction of the Bluetooth 4.0 version enables Bluetooth technology to exhibit lower power consumption in data transmission, as well as advantages such as low cost, low delay, and long effective connection distance, which promotes the development of Bluetooth in indoor localization.

The current mainstream fingerprint-based indoor localization approach includes two phases [12], namely offline and online phases. In the offline phase, site surveys are conducted and the regions of interest are

* Correspondence: jinxiaoxiaosx@163.com
[1]School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
[2]College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China

planned. Received Signal Strength (RSS) samples are collected at each planning Reference Point (RP) and saved to the location fingerprint database. In the online phase, by matching the RSS measured at the point to be positioned with the location fingerprint database, the location result of pedestrian is output. It can be seen that the process of constructing location fingerprint database in offline phase consumes a great deal of time and manpower, and is vulnerable to changes in the environment, such as the flow of crowd and Access Point (AP) change. When the area is too large, the cost required for constructing the fingerprint database cannot withstand. The above factors all limit the promotion and application of this technology. In this paper, we aim to construct a fingerprint database with significantly reduced workload as well as guarantee high localization accuracy in the online phase.

To achieve this goal, a new cost-efficient fingerprint database construction approach via dynamical collection approach and multi-quadric [13] RBF interpolation is proposed. In concrete terms, at fingerprints collection stage, we collect fingerprints on a few straight-line paths dynamically and then add error correction factors into pedestrian dead reckoning (PDR) part, to generate an accurate sparse fingerprint database. In addition, we apply multi-quadric RBF interpolation approach to interpolating each new fingerprints, and we also rely on GA [14] to find better shape parameter to improve interpolation accuracy.

All the experiments are conducted in a real-world indoor BLE environment to demonstrate the system practicability. The three main contributions of this paper are summarized as follows:

(a) The dynamic collection of fingerprint samples and PDR error correction factors are considered to construct an accurate raw sparse fingerprint database.

(b) The multi-quadric RBF interpolation algorithm is applied to expand the sparse fingerprint database, in which the impact of free shape parameter and cut-off radius on RSS estimation is studied also.

(c) The extensive experiments in a real-world indoor BLE environment demonstrate the effectiveness and efficiency of our system in terms of localization accuracy and fingerprint database construction effort.

The rest of the paper is organized as follows. Section 2 shows some related works concentrating on indoor Bluetooth localization and the way to reduce fingerprints calibration effort. In Section 3, we introduce the framework of the system, including using a dynamic collection of fingerprint samples to construct the sparse fingerprint database and using multi-quadric RBF to interpolate the sparse fingerprint database in offline phase. It also includes BLE localization in the online phase according to the

WKNN algorithm. Section 4 provides the experimental results, and finally, the conclusion of this paper is given in Section 5.

2 Related work

2.1 Indoor Bluetooth localization

Indoor Bluetooth localization approaches can be divided into three main categories, namely range detection, fingerprint matching, and propagation model-based approaches. The range detection-based approach makes use of the short distance transmission characteristics of the Bluetooth signal. When the user carries the mobile device into the range of signal coverage, the user's location can be perceived that achieves the range level localization accuracy. Anastasi and Chawathe et al. [15, 16] put forward the Cell-ID parameters to implement the Bluetooth localization system. In the localization system, the user's movement and standing can be determined to realize the room level localization. Signal intensity based approaches are mainly divided into two categories, namely fingerprint matching and propagation model-based approaches. Mo and Xiong et al. [17] propose to classify the region according to the proximity information of the signal to limit the search range. The other parameter based localization approaches mainly use link quality information, query feedback rate [18] and other parameters. Forno et al. [19] propose that the Bluetooth power level can be used to locate mobile users, and the surrounding Bluetooth devices are obtained by filtering different power levels to calculate the location of the user.

2.2 Collection effort reduction

To reduce the time and labor cost for fingerprint database construction, many studies have been paid significant attention to reducing the fingerprints collection efforts. In [20], according to the auto-building system, the authors construct the location fingerprint database. In [21], based on the signal propagation model, the authors estimate RSS at all new RPs. Ouyang et al. [22, 23] apply the generative and discriminative semi-supervised learning and use a few labeled RSS data and a large number of unlabeled RSS data to construct and enrich the fingerprint database. Liu et al. [24] propose to use the compressed sensing theory to construct a fingerprint database. Racko et al. [25] choose linear interpolation for the purpose of reducing the time needed for radio map creation and verify the approach in a corridor which is a narrow environment. Kubota et al. [26] propose an accurate interpolation for survey database, part of fingerprints are estimated using a path loss model containing wall attenuation. A novel approach to

adapt temporal radio map is proposed in [27] by offsetting the varying environment factors.

Different from the studies above, we firstly determine the capacity of fingerprint acquisition sample according to the change characteristics of the BLE signal and localization error of different percentile values of the fingerprint database. Then, the sparse fingerprint database is constructed by extracting dynamic sampling fingerprint samples and combining pedestrian movement parameters information. In addition, the multi-quadric RBF interpolation approach is applied to estimate RSS sequences of the new RPs, thus compensating the deficiency of sparse fingerprint database localization accuracy. The cut-off area radius and shape parameter in multi-quadric function are also studied and optimized, which improves the interpolation accuracy of sparse fingerprint database.

3 Methods

As shown in Fig. 1, the proposed fast construction approach of location fingerprint database in the offline phase, that is, using MEMS sensors to assist in the dynamic collection of BLE fingerprints and applying the RBF interpolation algorithm. First of all, using the MEMS sensors to measure the pedestrian motion parameters and calculating the location of fingerprint points, and then combining the BLE RSS, we construct the sparse fingerprint database. Second, the RBF interpolation algorithm is applied to enrich the sparse fingerprint database by interpolating each new RP from its neighbors. At the same time, in the online phase, the Weighted K Nearest Neighbor (WKNN) algorithm is

used to realize the accurate localization, which reduce the localization time.

3.1 Speed and heading angle estimation

According to the PDR algorithm [28], the pedestrian location is calculated from the current speed, heading angle, and location of a pedestrian at the last moment. In this paper, based on the output data of accelerometer, magnetometer, and gyroscope, the speed and heading angle of the pedestrian are obtained by combining speed estimation and heading angle calculation modules.

The speed and heading angle estimation process are as shown in Fig. 2. In speed estimation module, the gait detection is completed by using the periodic characteristics of the acceleration when the human body walks, and then the step length empirical model is applied to output the real-time step value. According to the correspondence between step length and step frequency and acceleration variance, the BP neural network is used to train model parameters. Finally, the speed is calculated based on the acceleration change form and step length information per second. In heading angle calculation module, according to the complementary characteristics of the three sensors, we use complementary filtering algorithm to optimize the output data of gyroscope and then update the quaternion based on Extended Kalman Filter (EKF) [29].

3.1.1 Speed estimation

The acceleration of pedestrians changes in the form of a sinusoidal wave during walking [30]. In this paper, the acceleration value of a pedestrian is collected by the
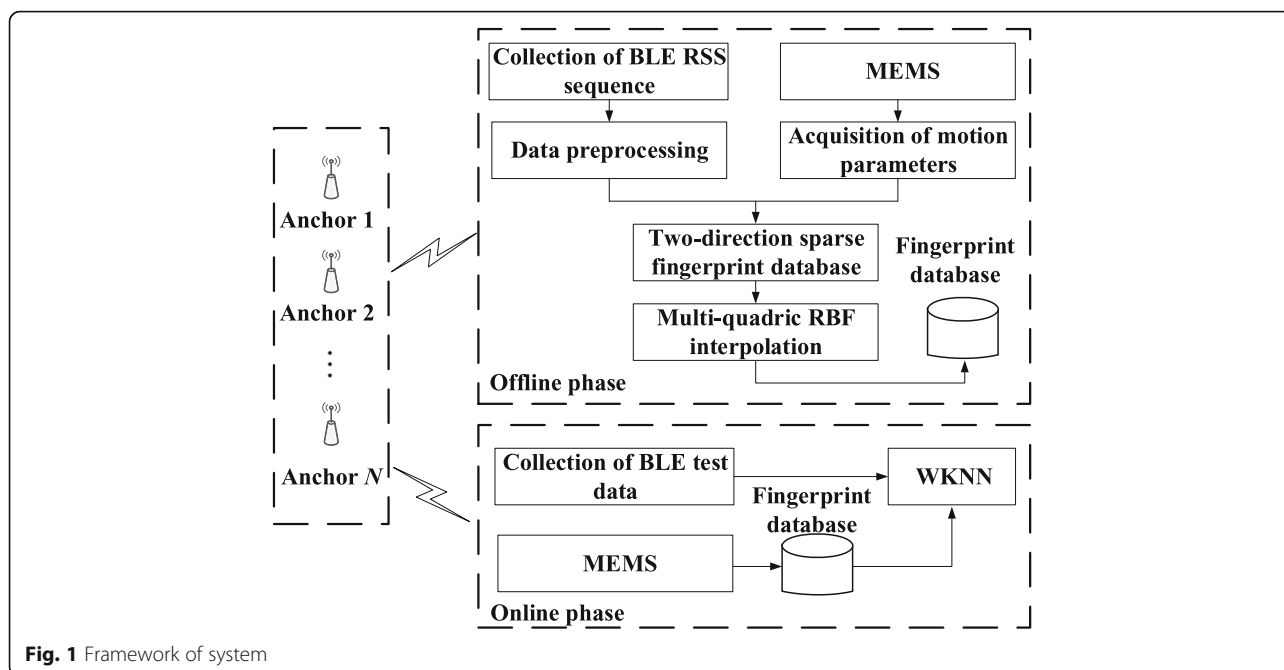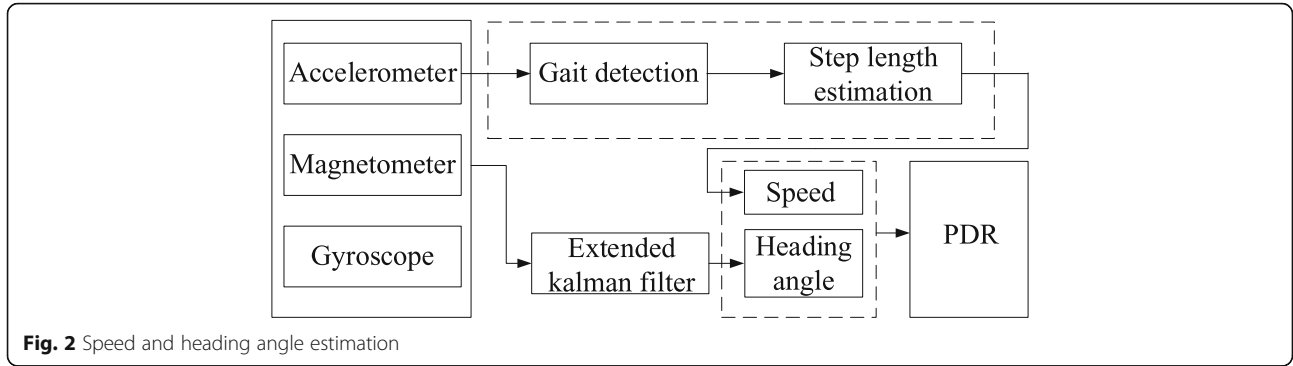

Fig. 1 Framework of system

Fig. 2 Speed and heading angle estimation

acceleration sensor, we calculate the modulus of three-axis acceleration to estimate the speed of pedestrian and detect the gate and steps of a pedestrian by using the periodic variation of acceleration during walking, in which acceleration modulus Acc_i^{total} is obtained by

$$Acc_i^{total} = \sqrt{\left(a_i^x\right)^2 + \left(a_i^y\right)^2 + \left(a_i^z\right)^2} \qquad (1)$$

where a_i^x, a_i^y, and a_i^z are acceleration at the i-th moment, in X axis, Y axis, and Z axis respectively.

Next, we estimate the pedestrian step length. Pedestrian step length is usually related to height, walking speed, and other factors [6]. To reasonably estimate the pedestrian step length, we select the nonlinear step length estimation model, which is an appropriate step length estimation model. Moreover, by using the calibration coefficient in the training model, the applicability and accuracy of the step length model are improved. The step length L_k can be expressed as

$$L_k = \rho \left(\sqrt{acc_{max}^k \text{-} acc_{min}^k} \right)^{1/4} \qquad (2)$$

where acc_{max}^k and acc_{min}^k are the maximum and minimum acceleration detected in k-th step. ρ is the calibration parameter, which is the ratio of true step length and estimated step size and can be obtained based on the neural network algorithm.

Finally, it is known that the sampling frequency of the sensor is f_s, the number of sampling points between two effective peaks is ΔN, and the step length of k-th step is L_k, we estimate the speed as

$$v_k = \frac{L_k f_s}{\Delta N} \qquad (3)$$

3.1.2 Heading angle
Based on the output data of accelerometer, magnetometer, and gyroscope, we calculate the attitude angle,

including the direction angle ϕ, roll angle γ, and pitch angle θ.

Assuming that the pedestrian is stationary and the handheld terminal is placed horizontally at the beginning, we initialize the filter. The output vector of three-axis accelerometer is expressed as

$$\begin{bmatrix} a_x & a_y & a_z \end{bmatrix}^T = C_n^b \begin{bmatrix} 0 & 0 & g \end{bmatrix}^T \qquad (4)$$

where $\begin{bmatrix} a_x & a_y & a_z \end{bmatrix}^T$ is the output vector of gravity acceleration in the carrier coordinate system, g is local acceleration and directional cosine matrix C_n^b is calculated by

$$C_n^b = \begin{pmatrix} \cos\gamma & 0 & \sin\gamma \\ \sin\theta \sin\gamma & \cos\theta & -\sin\theta \cos\gamma \\ -\cos\theta \sin\gamma & \sin\theta & \cos\theta \cos\gamma \end{pmatrix} \qquad (5)$$

Based on formulas (4) and (5), we have

$$\begin{cases} \gamma = \arctan \dfrac{a_x}{\sqrt{a_y^2 + a_z^2}} \\ \theta = \arctan \dfrac{-a_y}{a_z} \end{cases} \qquad (6)$$

Thus the three-axis magnetometer data is decomposed to horizontal direction as

$$\begin{cases} H_x = m_x \cos\theta + m_z \sin\theta \\ H_y = m_x \cos\gamma + m_y \sin\gamma \sin\theta \text{–} m_z \cos\theta \sin\gamma \end{cases} \qquad (7)$$

where $\begin{bmatrix} m_x & m_y & m_z \end{bmatrix}^T$ is the output vector of the magnetometer in the carrier coordinate system. H_x and H_y are the components of magnetic in the NES navigation coordinate system. Based on H_x and H_y, the heading angle can be calculated by

$$\varphi = \arctan \frac{H_x}{H_y} \qquad (8)$$

After the initialization process, next, we solve the real-time attitude angle and set the quaternion as the state variable of the filter, we have

$$X = [q_1 \ q_2 \ q_3 \ q_4]^T \tag{9}$$

In addition, in order to estimate the attitude angle more accurately, we select the output values of accelerometer and magnetometer to be as the observation, such that

$$Z = \begin{bmatrix} a_x & a_y & a_z & m_x & m_y & m_z \end{bmatrix}^T \tag{10}$$

According to the principle of the quaternion in the inertial navigation system, we define the state equation as

$$\dot{X} = \frac{1}{2} \begin{bmatrix} 0 & -(\omega_x - w_x) \\ (\omega_x - w_x) & 0 \\ (\omega_y - w_y) & -(\omega_z - w_z) \\ (\omega_z - w_z) & (\omega_y - w_y) \\ -(\omega_y - w_y) & -(\omega_z - w_z) \\ (\omega_z - w_z) & -(\omega_y - w_y) \\ 0 & (\omega_x - w_x) \\ -(\omega_x - w_x) & 0 \end{bmatrix} X \tag{11}$$

where ω_x, ω_y, and ω_z are the output of gyroscope and w_x, w_y, and w_z are measurement noise.

After the above data preparation, the optimal quaternion of the state variable is obtained by updating the EKF model. According to the conversion relationship [31] between quaternion and the Euler angle, the real-time heading angle is calculated as

$$\psi = \arctan \frac{-2(q_1 q_2 + q_0 q_3)}{(q_0^2 + q_1^2 - q_2^2 - q_3^2)} \tag{12}$$

3.2 Fingerprint database construction

In this paper, we use a dynamic fingerprint database construction algorithm based on MEMS sensors assisted BLE. The user holds the tag board containing MEMS sensors and BLE module, collects fingerprints on predetermined paths in the indoor area, and synchronizes the sensor data and BLE data of each moment by using time label. Then, the location information can be obtained by calculating pedestrian motion parameters such as speed and heading angle, and with the BLE RSS sequences

obtained at the RPs, a sparse location fingerprint database can be constructed.

Figure 3 shows the paths of fingerprint collection. For each path, the process is described as follows.

Step 1: According to the starting and ending point of the path L and the pre-set RP interval, the coordinates of standard location fingerprint database are obtained as

$$\begin{cases} X_i = x_0 + i^* L_{\text{step\_x}} \\ Y_i = y_0 + i^* L_{\text{step\_y}} \end{cases} \tag{13}$$

where $i \in (1, 2, \cdots, \frac{x_{\text{end}} - x_0}{L_{\text{step\_x}}})$, (X_i, Y_i) is the i-th RP coordinates, (x_0, y_0) and $(x_{\text{end}}, y_{\text{end}})$ are the starting point and ending point, and $L_{\text{step\_x}}$ and $L_{\text{step\_y}}$ are the pre-set RP interval.

Step 2: The disadvantage of MEMS localization algorithm is the cumulative error in the long-time period [28], which leads to large localization error. To address this problem, we add the error correction parameters to the PDR formula, such that

$$\begin{cases} x_j = x_0 + \sum_{n=1}^{j} vx_n + \varepsilon x_j \\ y_j = y_0 + \sum_{n=1}^{j} vy_n + \varepsilon y_j \end{cases} \tag{14}$$

where (x_j, y_j) is the location of a pedestrian at the j-th moment, and vx_j and vy_j are the two components of walking speed at the j-th moment, which is calculated by

$$\begin{cases} vx_j = v_j \sin(\text{head}_j) \\ vy_j = v_j \cos(\text{head}_j) \end{cases} \tag{15}$$

where head_j is heading angle at the j-th moment.

And correction factors εx_j and εy_j are defined as

$$\begin{cases} \varepsilon x_j = \frac{vx_j}{\sum_{n=1}^{\text{end}} vx_n} \left(x_{\text{end}} - L_{xpdr} \right) \\ \varepsilon y_j = \frac{vy_j}{\sum_{n=1}^{\text{end}} vy_n} \left(y_{\text{end}} - L_{ypdr} \right) \end{cases} \tag{16}$$

where $L_{xpdr} = \sum_{n=1}^{\text{end}} vx_n$ and $L_{ypdr} = \sum_{n=1}^{\text{end}} vy_n$ are the projection length of PDR result on X and Y axes.

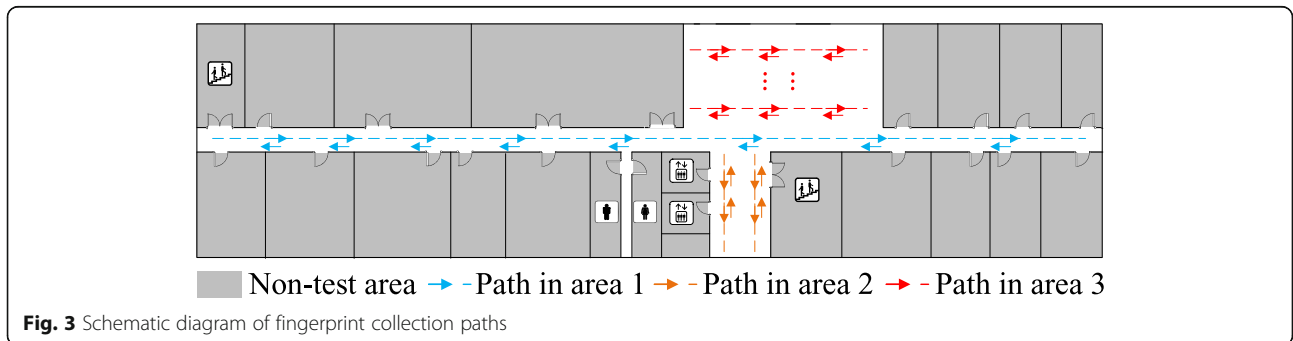Step 3: According to step 2, we obtain the coordinates of the pedestrian at each moment, and the

Fig. 3 Schematic diagram of fingerprint collection paths

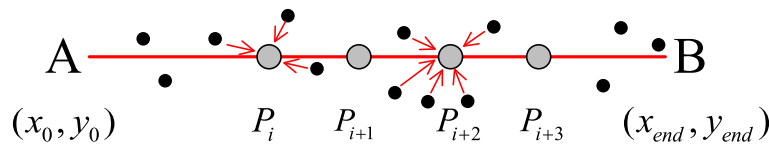Non-test area → -Path in area 1 → -Path in area 2 → -Path in area 3

Fig. 4 Fingerprint matching diagram

corresponding RSS sequence is also obtained. Next, we match the estimated pedestrian coordinates with standard fingerprint database coordinates according to the Nearest Neighbor (NN) algorithm [32].

As shown in Fig. 4, there are N locations near the path. For each location, calculating the Euclidean distance between the location and all standard coordinates, and matching to the corresponding standard coordinates in accordance with the NN algorithm. Taking the standard coordinates P_i as an example, the RSS is calculated by

$$\text{RSS}_n = \frac{\sum_{i=1}^{M} rssi_i}{M} \tag{17}$$

where RSS_n is the corresponding RSS on this point, and M is the total number of times that the point is stored in RSS sequence.

Step 4: The coarse location fingerprint database is generated in step 3, which is constructed by pedestrians location information and RSS at each moment. The walking speed of the pedestrian is about 1.3 m per second, so there are many fingerprint holes in each collection path, such as point P_{i+1}. To address this problem, the distance-based spatial fingerprints interpolation algorithm is selected, that is to calculate the mean RSS of adjacent RPs and assign it to RPs that without RSS.

In the process of constructing the database, we considered the shielding factors of the human body for BLE signal. At the same time, we consider the location fingerprints of different paths including the heading angle information, so we construct a multidirectional location fingerprint database. For example, if the heading angle of the path is 45° to 135°, the 90° direction database is set. Accordingly, 0°, 180°, and 270° direction databases are set up in turn. The RSS vector is expressed as

$$\psi_i^k = \left[\phi_{i,1}^k, \cdots, \phi_{i,j}^k, \cdots, \phi_{i,M}^k \right]^{\text{T}} \tag{18}$$

where $\phi_{i,j}^k (1 \le i \le P^k, 1 \le j \le M, 1 \le k \le 4)$ indicates the RSS value from j-th anchor received at the i-th RP. P^k is the number of RPs for the k-th direction database, and M is the number of anchor in the indoor area. The fingerprint database in the k direction can be expressed as

$$\Lambda^k = \begin{bmatrix} \psi_1^k & \psi_2^k & \cdots & \psi_{P^k}^k \\ x_1 & x_2 & \cdots & x_{P^k} \\ y_1 & y_2 & \cdots & y_{P^k} \end{bmatrix}^{\text{T}} \tag{19}$$

Because of the particularity of environment, the number of RPs in each direction fingerprint database is different, and the interpolation process behind is based on each direction database.

3.3 Fingerprint database interpolation

For convenience, we collect fingerprint data along the tile side line of the site and the length of a square tile is 0.8 m. We collect fingerprints by two tiles, that is, the particle size is $0.8 \times 1.6 \text{ m}^2$. Then, we continue to interpolate the sparse fingerprint database, so that the interval between the fingerprint database reference points is 0.8 m, that is, to get a smaller fingerprint database with smaller granularity.
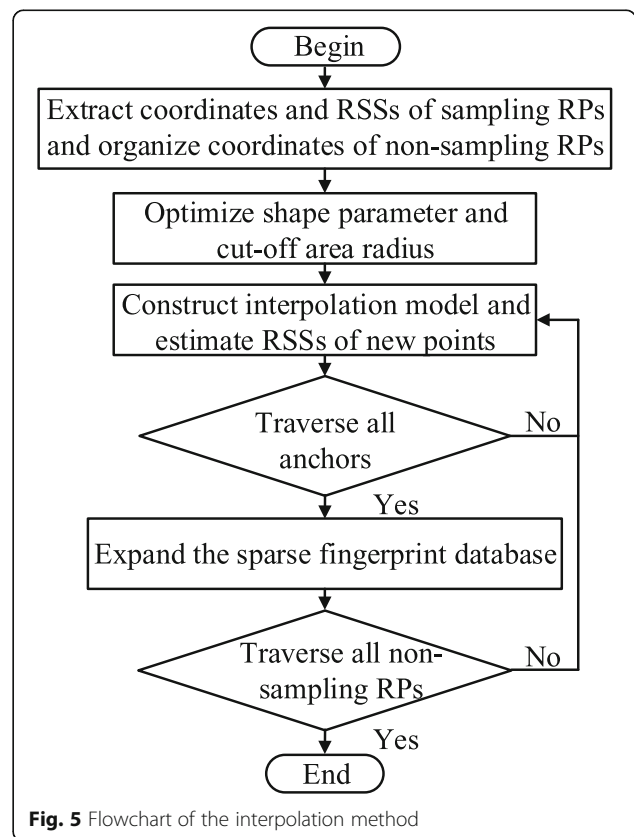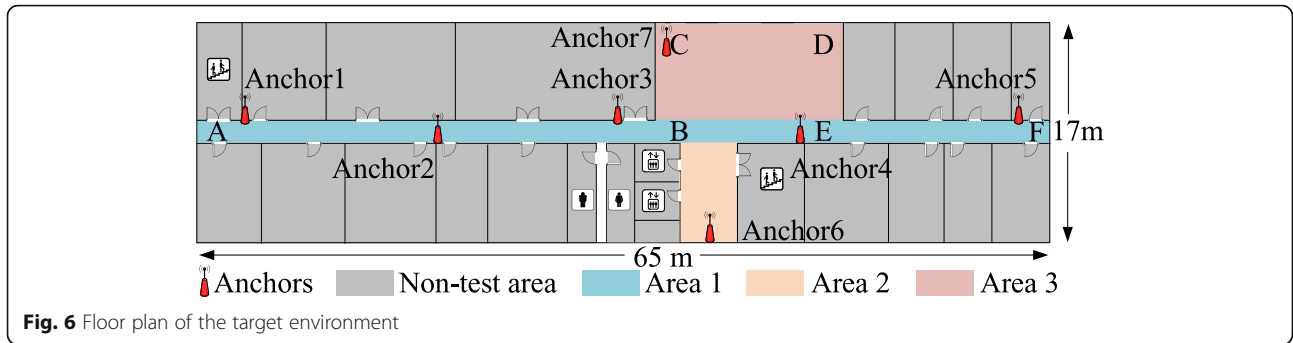

Fig. 5 Flowchart of the interpolation method

Fig. 6 Floor plan of the target environment

3.3.1 Interpolation model construction

The general function approximation approach mainly includes polynomial interpolation [33], Kriging model [34], neural network [35], and the RBF interpolation. The RBF approach receives extensive attention because of its advantages of less setting parameters and simple application. In this paper, the RBF interpolation approach is used to estimate RSS from the same anchor of RP in a given area. In addition, we especially study the impact of cut-off radius and shape parameter on interpolation precision. By optimizing the optimal values of the two parameters, the estimation error of new RPs RSS is reduced.

The radial basis function is defined as a monotonic function of the distance from any point x to center c in space, which is express as

$$\psi(x-c) = \varphi(\|x-c\|) \tag{20}$$

In this system, the Euclidean distance is applied to calculate the distance between two RPs. The function space composed of formula (20) and its linear combination can approximate any object function.

The approach of constructing the RBF interpolation model is described as follows. We set $\{x_j, y_j\}_{j=1}^{n} \in R^d \otimes R$

as a sampling point coordinate set, the RBF interpolation model is expressed as

$$f(x) = \sum_{j=1}^{N} \lambda_j \phi(\|x-x_j\|), x \in R^d \tag{21}$$

where λ_j is the weight coefficient of each sample point. $\sum_{j=1}^{N} \lambda_j \phi(\|x-x_j\|)$ is the base function.

Based on the coordinates set of sampling points, the base function matrix is expressed as

$$\phi_{n \times n} = \begin{pmatrix} \phi(\|x_1-x_1\|) & \phi(\|x_2-x_1\|) \\ \phi(\|x_1-x_2\|) & \phi(\|x_2-x_2\|) \\ \vdots & \vdots \\ \phi(\|x_1-x_n\|) & \phi(\|x_2-x_n\|) \\ \cdots & \phi(\|x_n-x_1\|) \\ \cdots & \phi(\|x_n-x_2\|) \\ \ddots & \vdots \\ \cdots & \phi(\|x_n-x_n\|) \end{pmatrix} \tag{22}$$

According to ϕ_{nxn} and the RSS sequence of the same anchor received at each sampling point, the weight coefficient matrix $(\lambda_1 \lambda_2 \cdots \lambda_n)$ can be calculated by
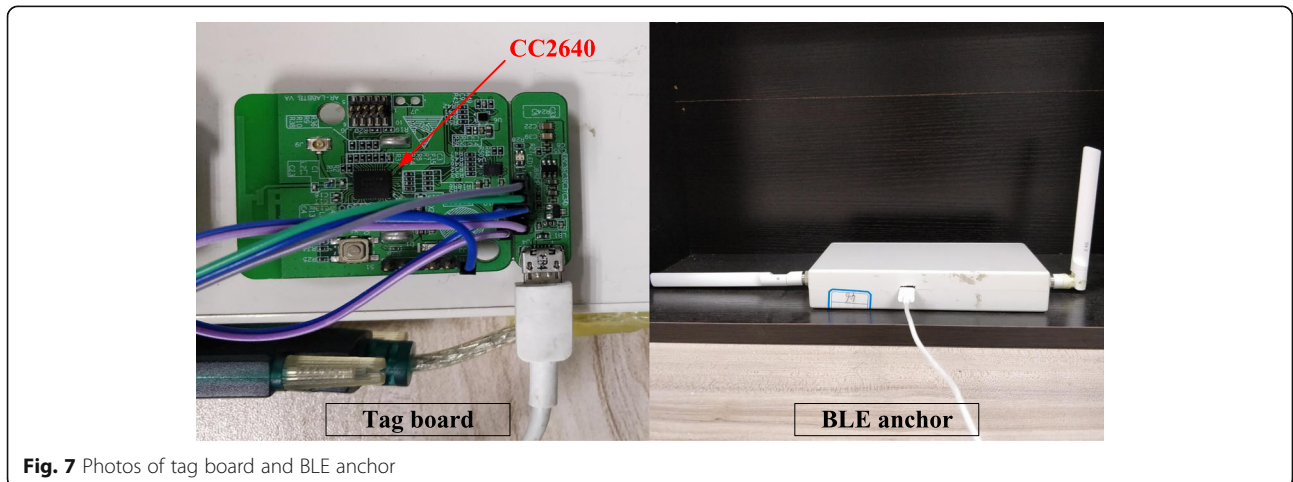

Fig. 7 Photos of tag board and BLE anchor

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_n \end{pmatrix}^T \begin{pmatrix} \phi(\|x_1-x_1\|) & \phi(\|x_2-x_1\|) \\ \phi(\|x_1-x_2\|) & \phi(\|x_2-x_2\|) \\ \vdots & \vdots \\ \phi(\|x_1-x_n\|) & \phi(\|x_2-x_n\|) \end{pmatrix}$$
$$\begin{matrix} \cdots & \phi(\|x_n-x_1\|) \\ \cdots & \phi(\|x_n-x_2\|) \\ \ddots & \vdots \\ \cdots & \phi(\|x_n-x_n\|) \end{matrix} = \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}^T \tag{23}$$

The condition for formula (23) to be solvable is that $\phi_{n \times n}$ is reversible. It has been proved that the above equations have unique solutions when the sampling points are different from each other [36]. Thus, formula (23) can be further expressed as

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_n \end{pmatrix}^T = \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}^T \begin{pmatrix} \phi(\|x_1-x_1\|) \\ \phi(\|x_1-x_2\|) \\ \vdots \\ \phi(\|x_1-x_n\|) \end{pmatrix}$$
$$\begin{matrix} \phi(\|x_2-x_1\|) & \cdots & \phi(\|x_n-x_1\|) \\ \phi(\|x_2-x_2\|) & \cdots & \phi(\|x_n-x_2\|) \\ \vdots & \ddots & \vdots \\ \phi(\|x_2-x_n\|) & \cdots & \phi(\|x_n-x_n\|) \end{matrix}^{-1} \tag{24}$$

After obtaining the weight coefficient matrix, we bring it into formula (21). By setting the coordinates of new RPs as input, the RSS values of the corresponding anchor can be solved.

3.3.2 Parameter optimization

The commonly used radial basis function includes Gauss distribution function of the Kriging method, multi-quadric function of Hardy, and the inverse multi-quadric function of Hardy. It has been proved in [36] that linear equation (23) have unique solvability when the kernel function is multi-quadric function, so multi-quadric function is used as kernel function in this paper. The base function formula can be expressed as

$$\phi(x) = \sqrt{c^2 + \|x\|^2} \tag{25}$$

where c is shape parameter, and $\|x\|$ is Euclidean distance between two points.

The shape parameter mainly affects the value of the base function. The authors in [37, 38] study the selection of shape parameter. However, these methods are based on the values of the sampling points, giving an empirical formula or a direct assignment. At the same time, the influence of condition number on the stability of the basis function matrix is not considered.

In this paper, we use the two criteria of root mean square error (RMSE) and the Pearson correlation coefficient P_{pear} [39] to select the optimized shape parameter value. For the selected new sampling points, we measure

the real RSS sequences and obtain the RSS sequences estimated by using multi-quadric RBF interpolation algorithm. The RMSE is the objective function, and condition number of matrix is used as a constraint condition, we obtain

$$\sqrt{\frac{\sum_{i=1}^{n} (y(x_i)-s(x_i,c))^2}{n}} = f(c) \tag{26}$$

where $y(x_i)$ is real measured RSS value of point x_i, $s(x_i,c)$ is the estimated RSS value of point x_i based on interpolation approach, n is the total number of selected sampling points, and c is shape parameter. The optimization process is expressed as formula (27), and the condition number is less than 10^{15} [40]. When the above conditions are satisfied, c corresponding to the max P_{pear} is selected.
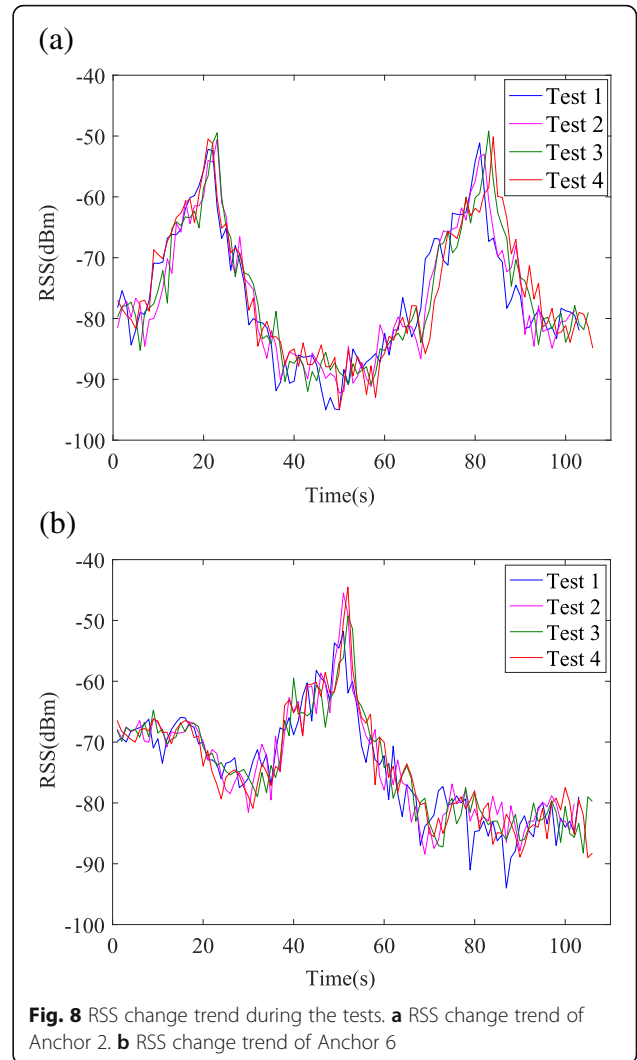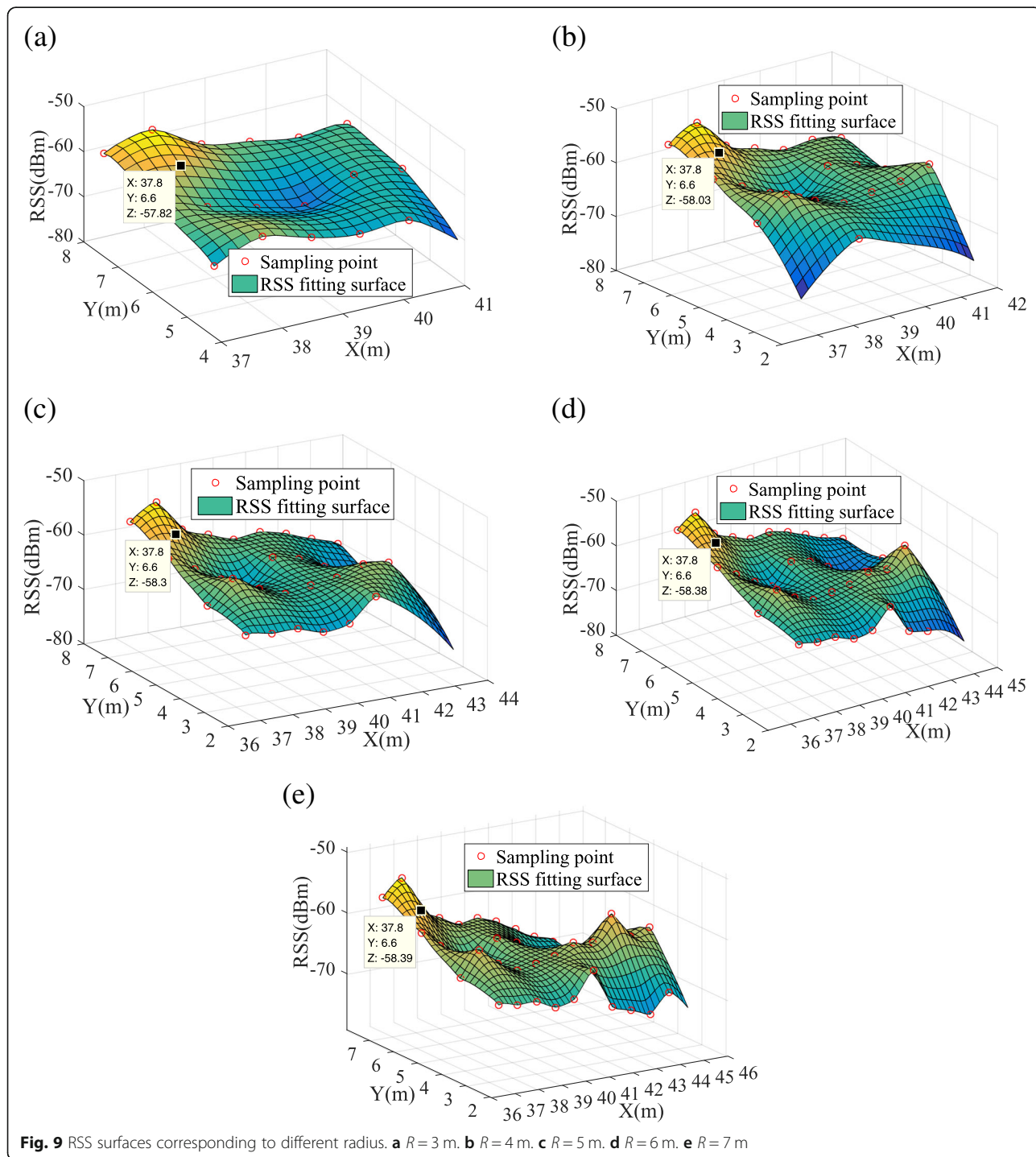

Fig. 8 RSS change trend during the tests. **a** RSS change trend of Anchor 2. **b** RSS change trend of Anchor 6

$$\begin{cases} \arg_c \min f(c) \\ \text{cond}(\phi) < 10^{15} \end{cases} \qquad (27)$$

3.3.3 Interpolation process description

In this paper, the optimized multi-quadric RBF is used to approximate the target function of RSS of sampling RPs, so as to achieve the purpose of

expanding the sparse fingerprint database, as shown in Fig. 5. We take the fingerprint samples in one direction as an example. First of all, coordinates and RSSs are extracted from sparse fingerprint database. Coordinates of non-sampling RPs are determined based on interpolation interval. Second, the optimal shape parameter is found through GA, and cut-off area radius is obtained

Fig. 9 RSS surfaces corresponding to different radius. **a** R = 3 m. **b** R = 4 m. **c** R = 5 m. **d** R = 6 m. **e** R = 7 m

based on the sampling fingerprints. Third, the circle cut-off area range is set with each non-sampling RP as a center. In the cut-off area, according to RBF interpolation approach, we construct an RSS surface of all sampling points from an anchor. Then we estimate RSSs of non-sampling RPs from the same anchor. By traversing all anchors, we can estimate the RSS sequence of a non-sampling RP. According to the above process, the RSS sequence of each non-sampling RP is estimated in turn. Coordinates of non-sampling RPs and the corresponding RSS sequences are combined stored into sparse fingerprint database.

4 Experimental results and analysis

4.1 Environmental layout

We select the first floor of an office building in school as the target environment (with demensions of 65 m by 17 m). As shown in Fig. 6, area 1, area 2, and area 3 are test areas, and the grey part is the non-test area. The target environment is a typical office environment composed of a corridor, a hall, and many indoor rooms. A total of 7 anchors are arranged in the environment, and the

Table 1 The difference between the estimated value and measured value of RSS (dBm)

| Coordinates	Radius (m)				
	3.0	4.0	5.0	6.0	7.0
(41.8, 6.6)	5.3	5.2	1.9	2.0	2.0
(42.6, 6.6)	3.1	1.6	1.9	2.0	2.0
(45.0, 6.6)	0.2	4.2	3.2	4.0	3.9
(45.8, 6.6)	0.5	1.9	2.2	2.3	2.4
(37.0, 5.0)	2.2	7.7	0.5	0.5	0.4
(37.8, 5.0)	5.3	7.8	8.8	8.8	8.8
(40.2, 5.0)	1.4	5.2	1.5	1.5	1.5
(41.0, 5.0)	0.9	0.6	0.5	0.6	0.6
(42.6, 5.0)	3.2	1.4	1.4	1.5	1.4
(43.4, 6.6)	4.1	5.2	4.2	5.3	5.2
(44.2, 6.6)	3.7	1.9	1.9	1.9	1.9
(46.6, 6.6)	0.5	5.0	1.0	1.0	1.0
(38.6, 3.4)	3.0	3.1	2.8	2.9	2.8
(39.4, 3.4)	3.8	2.4	2.5	2.5	2.6
(40.2, 3.4)	2.1	2.2	2.3	2.3	2.4
(41.8, 3.4)	1.7	5.9	1.2	1.2	1.2
(42.6, 3.4)	1.2	6.3	2.1	2.2	2.2
(43.4, 3.4)	1.9	1.6	1.8	1.8	1.9
(45.0, 3.4)	5.8	5.6	4.2	5.2	5.1
(45.8, 3.4)	1.3	8.2	0.8	0.8	0.7
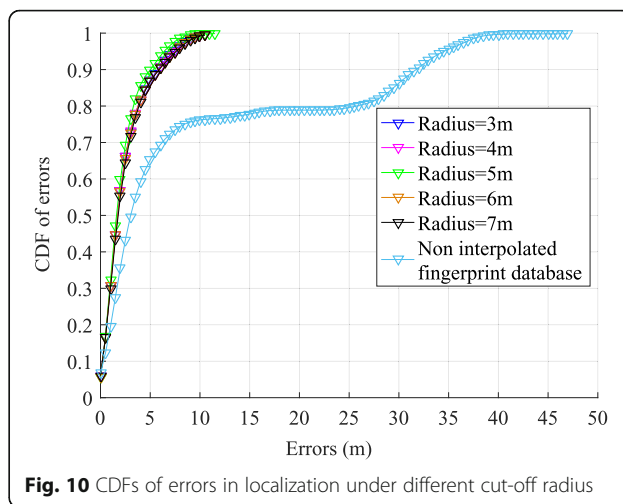
Fig. 10 CDFs of errors in localization under different cut-off radius

location of the anchor is shown in the location of the red small tower.

Our system consists of a portable tag board, several BLE anchors, a server, and a gateway. Tag board embedded CC2640 chip, including BLE module and sensor module. BLE anchor takes CC2640 as the core chip, including Bluetooth signal transmitting antenna. The server is a computer. The photos of the tag board and BLE anchor are shown in Fig. 7.

4.2 Parameter determination

4.2.1 Fingerprint sample capacity

There are seven anchors in the test environment, of which Anchor 2 is in the corridor and Anchor 6 is in the lobby. With Anchor 2 and Anchor 6 as representatives, Fig. 8a and b show changes of pedestrian receiving the RSS of the Anchor 2 and Anchor 6 respectively. From this figure, we can find that the BLE fingerprints collected dynamically by BLE platform have more stable characteristics, besides the signal jitter. Therefore, we
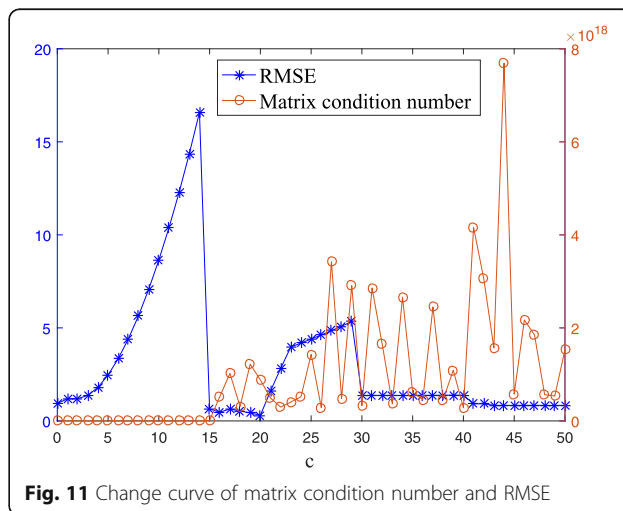

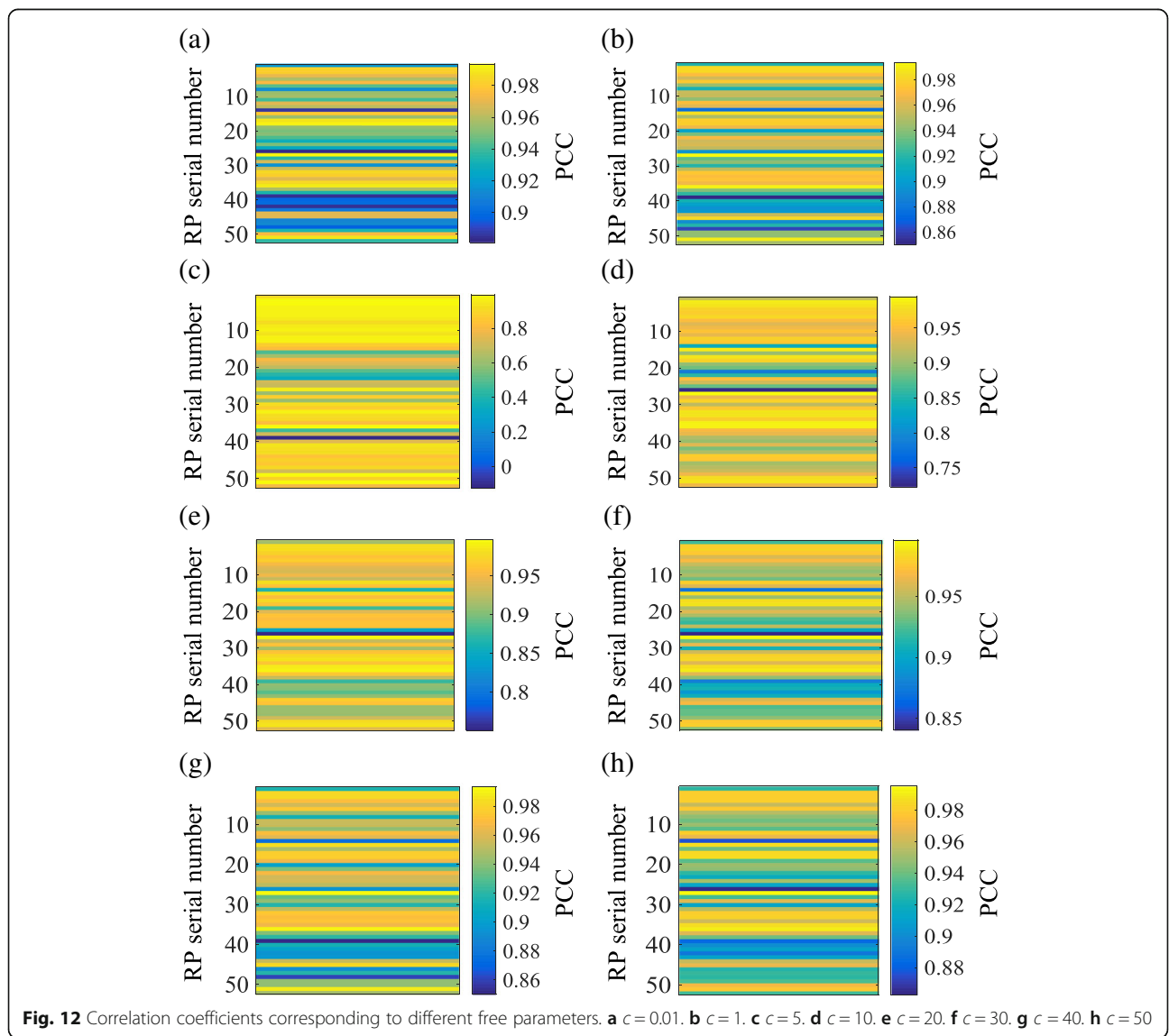**Fig. 11** Change curve of matrix condition number and RMSE

only walk one way back and forth on the pre-determined straight path, that is, collecting two direction fingerprint sequences on each path.

4.2.2 Cut-off radius

The multi-quadric RBF interpolation approach is selected to estimate RSS values of new sampling points. To verify the validity of this approach, we randomly select a sampling point that the coordinates is (37.8, 6.6) and the measured RSS is − 61.09 dBm. Based on the coordinates of sampling points and RSS of Anchor 7, we obtain RSS surface. Figure 9 shows RSS surfaces corresponding to radius of 3 m, 4 m, ···, and 7 m. From this figure, the estimated RSS values for the reference point are − 57.82 dBm, − 58.03 dBm, − 58.3 dBm, − 58.38 dBm, and − 58.49 dBm respectively. Therefore, it is feasible to

estimate the RSS value of the new RPs by using RBF interpolation algorithm within a certain range.

In order to select the cut-off area radius suitable for estimating RSS values for most new RPs, we randomly selected 20 test points to estimate the difference between the RSS values and measured RSS values of these RPs, as shown in Table 1. It can be seen that when the radius is 3 m and 4 m, there exist large errors, and when the radius is 5 m, 6 m, and 7 m, the estimation errors are similar and the whole estimated errors are smaller. In addition, through the statistical analysis of localization results of several sets of test data, we get the CDFs of localization errors, as shown in Fig. 10. We can find from this figure that the interpolated fingerprint database is superior to the non interpolated fingerprint database in localization performance, so it is necessary to interpolate the sparse fingerprint database. And when the cut-off radius is 5 m, the
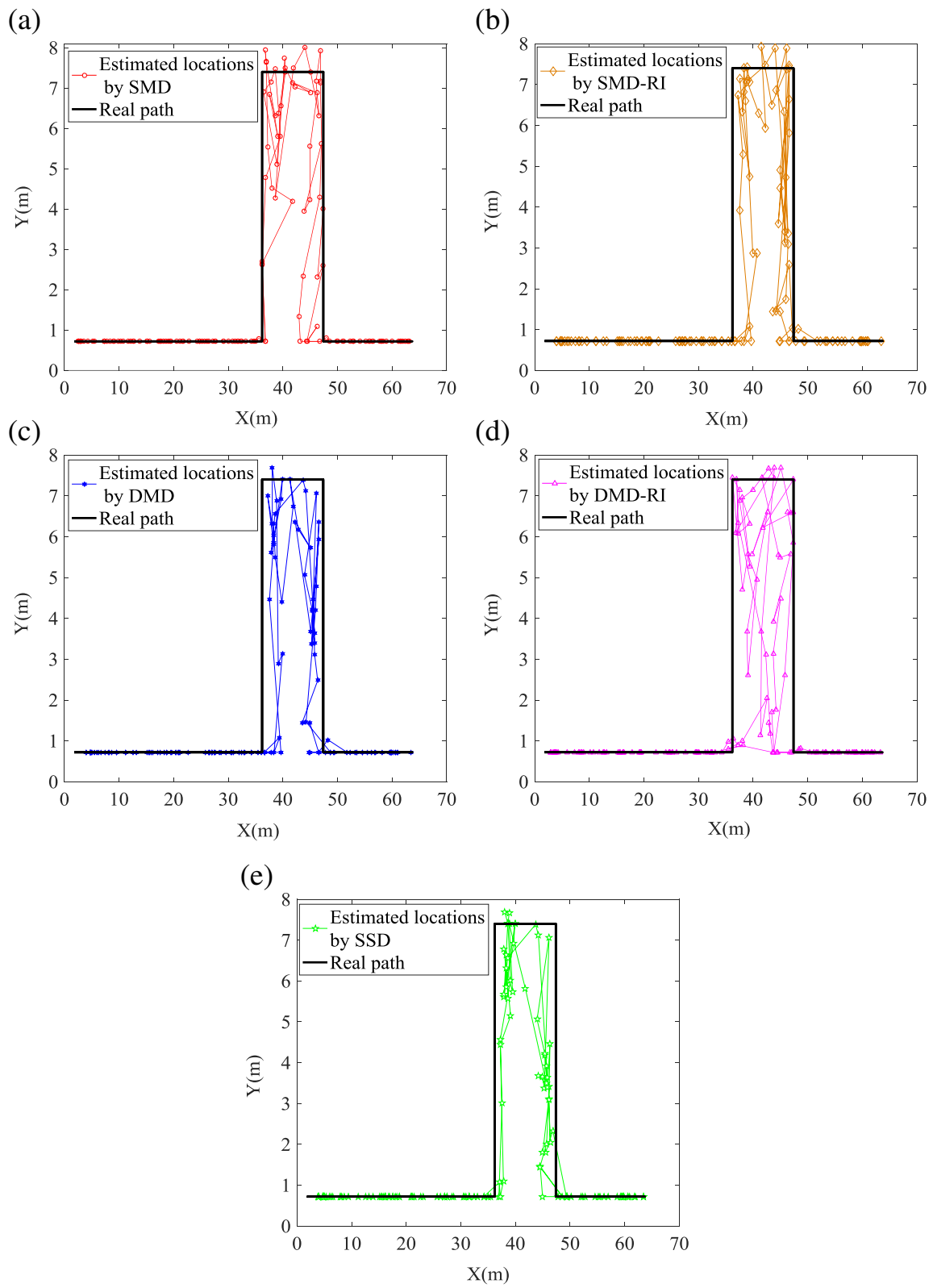

Fig. 12 Correlation coefficients corresponding to different free parameters. **a** $c = 0.01$. **b** $c = 1$. **c** $c = 5$. **d** $c = 10$. **e** $c = 20$. **f** $c = 30$. **g** $c = 40$. **h** $c = 50$

Fig. 13 Estimated locations by using different types of fingerprint database

localization accuracy performs best, the probability of errors within 3 m is 76.58%.

Based on the above experimental results, we set the cut-off area radius as 5 m.

4.2.3 Free shape parameter

Figure 11 shows the relationship between the matrix condition number of base function matrix and shape parameter and the relationship between RMSE of RSS estimation and shape parameter. It can be seen that when the range of changes in c is 0.01 to 50, RMSE increases with the increase of c, and oscillates after c is 14. The matrix condition number is more than 10^{15} when c is 14. Figure 12 shows that when c equals to different values, the Pearson correlation coefficient between the estimated RSS values and the measured RSS values of 52 RPs. It can be seen that when c is 0.01, the overall correlation coefficient is larger, that is, the estimated RSSs is closer to the real RSSs. In addition, in order to determine the specific shape parameter, we determine that the range of c is from 0.001 to 5 and we select a genetic algorithm to optimize the shape parameter. By setting the popsize 50, crossover probability 0.8, mutation probability 0.04, and number of iterations 200, the output shape parameter is set to be 0.07.
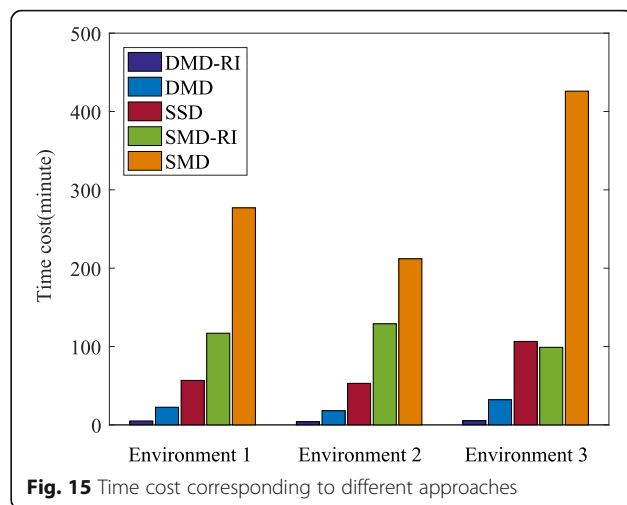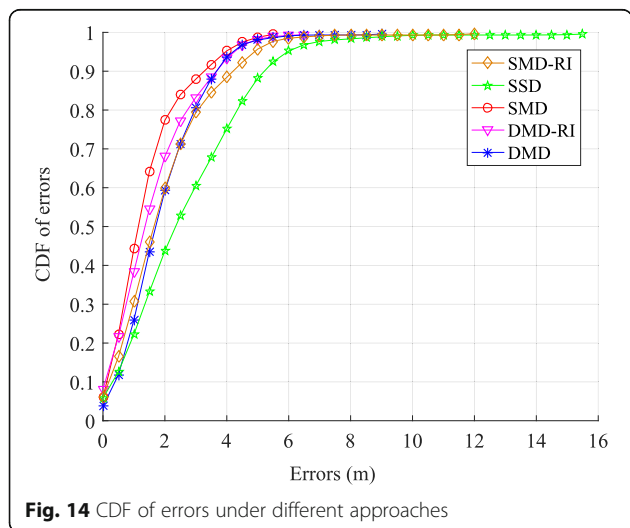
4.2.4 Localization result

In the target environment, we complete the collection work of Static Single Direction (SSD) database, Static Multi Direction (SMD) database, Dynamic Multi Direction (DMD) database, Static Multi Direction based RBF Interpolated (SMD-RI) database, and Dynamic Multi Direction based RBF Interpolated (DMD-RI) database proposed in this paper and analyze the localization performance of the 5 fingerprint database through dynamic test data. The test path starts from point A, along

Anchor 3 to point B, then to point C, point D, and point E in turn, along the corridor to point F, and finally back to point A along the original path. Figure 13 shows the estimated locations using several fingerprint databases. To further prove the result, the corresponding CDFs of errors are shown in Fig. 14. From this figure, we can find that the proposed fingerprint database is superior to DMD, SMD-RI, and SSD fingerprint databases and slightly poorer than SMD fingerprint database in terms of localization performance.

4.2.5 Time cost for fingerprints collection

To compare the time cost of the five kinds of fingerprint database construction approaches, we select 3 indoor environments with different size, i.e., environment 1 with $256\,\mathrm{m}^2$, environment 2 with $374\,\mathrm{m}^2$, and environment 3 with $288\,\mathrm{m}^2$. As shown in Fig. 15, it can be seen that the time cost for DMD-RI approach that proposed in this paper is less than that of other approaches. By taking environment 1 as an example, the proposed DMD-RI approach requires collecting two direction RSS samples on 5 paths in the hall and 1 long path in the corridor, which cost $290(=5 \times 16 \times 2 + 65 \times 2)$ s. In DMD approach, the time cost is $1350(=10 \times 3 \times 16 \times 2 + 65 \times 2 \times 3)$ s, in which two direction RSS samples on 10 paths in the hall and 1 long path in the corridor are collected 3 times. In the SSD approach, the time cost is $4160(=208 \times 20)$ s, in which 20 RSS samples at each of the 88 labeled RPs are collected. In SMD-RI approach, the time cost is $7040(=88 \times 20 \times 4)$ s, in which 20 RSS samples in 4 directions at each of the 88 labeled RPs are collected. At last, the time cost for SMD fingerprint database construction is $16640(=208 \times 4 \times 20)$ s, in which 20 RSS samples in 4 directions at each of the 208 labeled RPs are collected (Fig. 15).

Fig. 14 CDF of errors under different approaches

Fig. 15 Time cost corresponding to different approaches

5 Conclusion

In this paper, we have proposed a new approach of reducing the effort of fingerprint database construction for indoor BLE localization by decreasing RSS sample capacity, collecting the fingerprints dynamically and interpolating raw sparse fingerprint database by the proposed DMD-RI approach. In the process of constructing a sparse fingerprint database, the PDR result is more accurate by adding error correction factors, and the accuracy of directional fingerprint database is better by adding the heading angle. In the interpolation process, we first optimize shape parameter and cut-off area radius and then use multi-quadric RBF approach to estimate RSSs of new RPs, which makes the interpolation accuracy higher. Extensive experimental results show that the proposed approach is able to reduce much time cost for fingerprints collection while maintaining high localization accuracy. In the future, we will study the construction of an accurate multi-directional fingerprint database by extracting accurate crowdsourcing data.

Abbreviations
AP: Access Point; BLE: Bluetooth Low Energy; DMD: Dynamic Multi Direction; DMD-RI: Dynamic Multi Direction based RBF Interpolated database; EKF: Extended Kalman Filter; GA: Genetic Algorithm; GPS: Global Positioning System; MEMS: Micro Electro Mechanical Systems; NN: Nearest Neighbor; PDR: Pedestrian Dead Reckoning; RBF: Radial Basis Function; RFID: Radio-frequency identification; RMSE: Root Mean Square Error; RP: Reference Point; RSS: Received Signal Strength; SMD: Static Multi Direction; SMD-RI: Static Multi Direction based RBF Interpolated database; SSD: Static Single Direction; UWB: Ultra Wideband; WKNN: Weighted K Nearest Neighbor; WLAN: Wireless local area network

Availability of data and materials
Not applicable.

Authors' contributions
The authors have contributed jointly to all parts on the preparation of this manuscript, and all authors read and approved the final manuscript.

Authors' information
Not applicable.

Competing interests
The authors declare that they have no competing interests.

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

References
1. M. Jia, Q.G.Z. Yin, Design for spectrum efficient IoT network. IEEE Internet Things J. (2017), pp. 1–1
2. M. Jia, Q. Guo, X. Gu, Broadband hybrid satellite-terrestrial communication systems based on cognitive radio toward 5G. IEEE Internet Things J. **23**(6), 96–106 (2016)
3. M. Jia, X. Gu, X. Liu, Joint cooperative spectrum sensing and channel selection optimization for satellite communication systems based on cognitive radio. Int. J. Satell. Commun. Netw. **35**(2), 139–150 (2017)
4. M. Jia, Z. Yin, X. Liu, Joint cooperative spectrum sensing and spectrum opportunity for satellite cluster communication networks. Ad Hoc Netw. **58**(C), 231–238 (2016)
5. V. Knoop, P. Bakker, Lane determination with GPS precise point positioning. IEEE Trans. Intell. Transp. Syst. **33**(99), 1–11 (2017)
6. A. Khalajmehrabadi, N. Gatsis, Structured group sparsity: a novel indoor WLAN localization, outlier detection, and radio map interpolation scheme. IEEE Trans. Veh. Technol. **66**(7), 6498–6510 (2017)
7. R. Faragher, R. Harle, Location fingerprinting with bluetooth low energy beacons. IEEE J. Sel. Areas Commun. **33**(11), 2418–2428 (2015)
8. L. Segers, A. Braeken, J. Tiete, Ultrasonic multiple-access ranging system using spread spectrum and mems technology for indoor localization. Sensors **14**(2), 3172–3187 (2014)
9. C. Wu, R. Chang, S. Cheng, A data filtering strategy using cluster architecture in radio frequency identification system. Int. J. Radio Freq. Identif. Technol. Appl. **4**(2), 149–161 (2013)
10. E. Bogdani, D. Vouyioukas, N. Nomikos, Localization error modeling of hybrid fingerprint-based techniques for indoor ultra-wideband systems. Telecommun. Syst. **63**(2), 223–241 (2016)
11. Z. Jianyong, L. Haiyong, C. Zili, in *International Conference on Indoor Positioning and Indoor Navigation*. RSSI based Bluetooth low energy indoor positioning (2015), pp. 526–533
12. A. Abusara, M.S. Hassan, M.H. Ismail, Reduced complexity fingerprinting in WLAN-based indoor positioning. Telecommun. Syst. **65**(3), 407–417 (2017)
13. A. Saranli, B. Baykal, Complexity reduction in radial basis function (RBF) networks by using radial b-spline functions. Neurocomputing **12**(1), 183–194 (1998)
14. S. He, B. Ji, S.H. Chan, Chameleon: survey-free updating of a fingerprint database for indoor localization. Leonardo J. Sci. **3**(4), 44–58 (2004)
15. G. Anastasi, R. Bandelloni, M. Conti, *Experimenting an indoor bluetooth-based positioning service*, International Conference on Distributed Computing Systems Workshops (2003), pp. 480–483
16. S.S. Chawathe, *Beacon placement for indoor localization using bluetooth*, International IEEE Conference on Intelligent Transportation Systems (2008), pp. 980–985
17. R. Momose, T. Nitta, M. Yanagisawa, *An accurate indoor positioning algorithm using particle filter based on the proximity of bluetooth beacons*, IEEE, Global Conference on Consumer Electronics (2017), pp. 1–5
18. M.S. Bargh, R. de Groote, *Indoor localization based on response rate of bluetooth inquiries*, ACM International Workshop on Mobile Entity Localization and Tracking in Gps-Less Environments (2008), pp. 49–54
19. F. Forno, G. Malnati, G. Portelli, Design and implementation of a bluetooth ad hoc network for indoor positioning. IEEE proceedings-Software **152**(5), 223–228 (2005)
20. L. Ma, Y. Fan, Y. Xu, *Pedestrian dead reckoning trajectory matching method for radio map crowdsourcing building in wifi indoor positioning system*, IEEE International Conference on Communications (2017), pp. 1–6
21. D.P. Widyawan, M. Klepal, *Influence of predicted and measured fingerprint on the accuracy of rssi-based indoor location systems*, The Workshop on Positioning (2017), pp. 145–151
22. R.W. Ouyang, A.K. Wong, C.T. Lea, Indoor location estimation with reduced calibration exploiting unlabeled data via hybrid generative/discriminative learning. IEEE Trans. Mob. Comput. **11**(11), 1613–1626 (2012)
23. Y. Tu, Y. Lin, J. Wang, Semi-supervised learning with generative adversarial networks on digital signal modulation classification. Comput Mater Continua **55**(2), 243–254 (2018)
24. J.L. Liu, Y.H. Wan, B.G. Xu, *A novel indoor positioning method based on location fingerprinting*, International Conference on Communications, Circuits and Systems, vol 2 (2013), pp. 239–242

25. J. Racko, J. Machaj, P. Brida, Wi-fi fingerprint radio map creation by using interpolation. Procedia Eng **192**, 753–758 (2017)
26. R. Kubota, S. Tagashira, Y. Arakawa, *Efficient survey database construction using location fingerprinting interpolation*, International Conference on Advanced Information NETWORKING and Applications IEEE, vol.3 (2013), pp. 469–476
27. J. Yin, Q. Yang, L. Ni, *Adaptive temporal radio maps for indoor location estimation*, IEEE International Conference on Pervasive Computing and Communications (2005), pp. 85–94
28. G. Chen, X. Meng, Y. Wang, Integrated WiFi/PDR/smartphone using an unscented Kalman filter algorithm for 3d indoor localization. Sensors **15**(9), 24595–24614 (2015)
29. R. Toledo-Moreo, D. Gruyer, A. Lambert, *A theoretical analysis of the extended Kalman filter for data fusion in vehicular positioning* (International Conference on ITS Telecommunications, 2011)
30. D. Uritani, C. Sakamoto, T. Fukumoto, Effect of floating toes on knee and trunk acceleration during walking: a preliminary study. J. Phys. Ther. Sci. **29**(2), 361–364 (2017)
31. S. Han, J. Wang, A novel method to integrate IMU and magnetometers in attitude and heading reference systems. J. Navig. **64**(4), 727–738 (2001)
32. N. Wang, A.E. Raftery, Nearest-neighbor variance estimation. Publ. Am. Stat. Assoc. **97**(460), 994–1019 (2002)
33. C. Fassino, H.M. Möller, Multivariate polynomial interpolation with perturbed data. Numer. Algorithms **71**(2), 273–292 (2016)
34. S.S. Jan, S.J. Yeh, Y.W. Liu, Received signal strength database interpolation by kriging for a Wi-Fi indoor positioning system. Sensors **15**(9), 21377–21393 (2015)
35. V. Nourani, A. Babakhani, Integration of artificial neural networks with radial basis function interpolation in earthfill dam seepage modeling. J. Comput. Civ. Eng. **27**(2), 183–195 (2013)
36. J.H. Jung, V.R. Durante, An iterative adaptive multiquadric radial basis function method for the detection of local jump discontinuities. Appl. Numer. Math. **59**(7), 1449–1466 (2009)
37. E.J. Kansa, R.E. Carlson, Improved accuracy of multiquadric interpolation using variable shape parameters. Comp. Math. Appl **24**(12), 99–120 (1992)
38. S. Rippa, An algorithm for selecting a good value for the parameter c, in radial basis function interpolation. Adv. Comput. Math. **11**(2–3), 193–210 (1999)
39. H. Zhou, Z. Deng, A new sampling method in particle filter based on Pearson correlation coefficient. Neurocomputing **216**, 208–215 (2016)
40. W. Yue-xing, X. Lin, C. Xiao-qian, *The radial basis function shape parameter chosen and its application in engneering*, IEEE International Conference on Intelligent Computing and Intelligent Systems (2009), pp. 79–83