

RESEARCH

Open Access



Detection of an unknown radio transmitter using an enhanced K -nearest neighbor algorithm based on virtual reference point and RSSD information

Liyang Zhang, Taihang Du* and Chundong Jiang

Abstract

Accurate detection of the unknown radio transmitter (URT) is crucial to combat illegal occupation of radio signal resources and protect communication system from harmful signal interference. The fingerprint positioning technique based on received signal strength (RSS) is famous for requiring no extra equipment, antenna arrays, and time synchronization. However, conventional RSS-based fingerprint positioning techniques that utilize K -nearest neighbor (KNN) method are confronted with problems when the positioning target is radio transmitter with unknown emission strength and frequency. Moreover, they not only cannot realize the precise localization of the URT but also only rely on pre-set reference points in the fingerprint database. In this paper, a new KNN-based geo-location approach using received signal strength difference (RSSD) information and virtual reference point is proposed to estimate an URT location. To obtain more accurate RSSD measurements, a RSSD-based filtering method by calculating the Euclidean distance between each sampling RSSD and the average value is devised to establish the fingerprint database. To achieve higher positioning accuracy, we combine KNN technique with the virtual reference (VR) point to propose RSSD-VRKNN algorithm. The simulation results show that the proposed scheme can obtain the best positioning performance compared with the conventional KNN and weighted K -nearest neighbor (WKNN) techniques. The performance and feasibility of our proposed algorithm are verified through extensive experiments.

Keywords: Radio transmitter detection, Received signal strength difference (RSSD), Position location, Virtual reference (VR) point, K -nearest neighbor (KNN)

1 Introduction

Position location (PL) service has received great attention recently with the development of science and technology, and PL service has been applied in many important fields and occasions such as location, disaster relief, security, navigation, and tracking [1]. Currently, complex and various kinds of radio signals emitted by radio transmitters have been widely used and existed in our daily life. Most communication systems used in wireless applications, such as mobile searching, emergency communications, location-based interference management, and public safety, rely on radio signals and are susceptible to

interference by illegal signals [2, 3]. Therefore, accurate detection of the radio transmitter, as the focus of considerable research efforts, has received significant recent interest, and it is crucial to strengthen management of radio signals and realize the reasonable use of radio spectrum resources. However, most of the previous research works are more focus on the signal receiver positioning. The precise localization of an unknown radio transmitter (URT) is considered to be more challenging than the conventional issue of positioning in wireless sensor networks because the important parameters such as the emission strength and frequency of the URT are unknown.

Many kinds of measurements can be used as the positioning characteristic parameters, such as time of arrival (ToA) [4], time difference of arrival (TDoA) [5], frequency different of arrival (FDoA) [6], received signal strength

*Correspondence: thdu@hebut.edu.cn
School of Artificial Intelligence, Hebei University of Technology, Hongqiao,
Tianjin 300130, China

(RSS) [7], angle of arrival (AoA) [8, 9], and their combination [10–12]. Conventional positioning techniques rely on these measurements, and each has its own merits and drawbacks. The ToA-based positioning algorithms rely on measurements of signal travel times between the radio transmitter and access point (AP). The TDoA-based technique can be employed with no synchronization between the radio transmitter and AP. However, the reference nodes are needed to be synchronized in time. Both ToA-based and TDoA-based methods are only suitable for line-of-sight (LoS) positioning scenario. In comparison, AoA-based algorithms utilize the realtime measuring angles of the AP from the radio transmitter without requiring appropriate synchronization, emitter power, and time stamp. Nevertheless, due to the influences of none line-of-sight (NLoS) positioning scenario and multi-path, the positioning accuracy of AoA-based technique will be greatly affected. Among all of them, RSS-based positioning technique is more popular with the cost-effective. There is no need to prepare an antenna array or meet time synchronization requirement, because the RSS measurement contains the resultant information.

Since there are many kinds of measurement parameters as described above, a lot of positioning methods have been developed. Geometric positioning techniques calculate the location of the positioning target as the intersection of position line, which is between the positioning target and AP. Triangulation and trilateration are typical geometric positioning methods, and the RSS, AoA, ToA, and TDoA are commonly used as measurement parameters. In addition, some hybrid systems such as TDoA-AoA [13] have been employed in geometric positioning method. One of the disadvantages of geometric positioning techniques is that there is no theoretical process about the noise interference. Unlike the geometric positioning techniques, the statistical positioning approach provides a theoretical framework that contains the noise factor. The main idea is to assume that the measurement parameter distributions are similar to Gaussian probability distributions. Bayesian [14] and maximum likelihood (ML) [15] estimation are the typical statistical positioning methods. Among all positioning techniques, the RSS-based fingerprint positioning technique is popular with the advantages of no additional hardware, low cost, and wide positioning range [16]. Fingerprint positioning method consists of two parts: off-line database establishment phase and on-line positioning phase. The location information collected from the positioning area is stored in the fingerprint database, which includes the spatial coordinates of reference points and RSS information. After collecting the fingerprint information, the location of positioning target needs to be estimated. While the real-time data from the target is recorded, the estimated location can be obtained

by the appropriate algorithm. However, the received RSS sensing data may contain errors and affects the positioning precision due to various abnormal conditions, such as device failure and malicious cases. A novel robust relative fingerprinting-based passive localization algorithm via a data cleansing approach was proposed in [17]. In addition, to establish and maintain RSS-based fingerprint database more conveniently, Zuo et al. [18] proposed a multi-phase fingerprint map localization method based on interpolation.

1.1 Related work

Currently, the most widely adopted fingerprint localization algorithms are the K -nearest neighbor (KNN) [19] algorithm and weighted KNN (WKNN) [20] algorithm due to the low complexity suitable for practical application. Taking an example of KNN, it uses the on-line RSS to search for K smallest Euclidean distance of known locations from the fingerprint database by root mean square errors principle. The estimated position is obtained by averaging the locations of the selected K reference points from the database. If the distance in signal space is used as weighted average, higher positioning accuracy can be obtained on the basis of KNN. A soft range limited K -nearest neighbor (SRL-KNN) localization fingerprinting algorithm was proposed in [21], which scales the fingerprint distance by a range factor related to the physical distance between the user's previous position and the reference location in the database to reduce the spatial ambiguity in localization. Nevertheless, due to the fact that equal RSS differences at different RSS levels may not mean equal differences in geometrical distances in the calculation of signal distances between different RSS vectors, Li et al. [22] proposed a feature-scaling-based k -nearest neighbor (FS-kNN) algorithm for achieving improved localization accuracy. Since the accuracy of KNN-based algorithm using Euclidean distance is not high enough due to the ignorance of statistical regularities from the training set, an improved method combining the Manhattan distance with the WKNN algorithm was proposed to distinguish the influence of different reference nodes [23]. To obtain the optimized node location estimate, Fang et al. [24] proposed an optimal WKNN (OWKNN) algorithm for wireless sensor network (WSN) fingerprint localization in a noisy environment. In order to eliminate incorrect neighboring reference points and to avoid selected reference points located only on one side of the test point, an improved neighboring reference points selection method [25] was proposed based on their physical distances to the test point, instead of the widely used positions of the reference points.

However, whether KNN-based algorithm or WKNN-based algorithm, only the reference points in the database are selected. Since the location information of reference

point which is closer to the actual target is expected, we propose an improved KNN algorithm combined with the virtual reference point. In addition, the conventional RSS-based fingerprint positioning technology is generally used to solve the localization problem of signal receiver. On the contrary, the positioning target in our research is an URT, and it is a process of using signal receiver to find the transmitter. Due to the unknown information parameters such as emission strength and frequency of the radio transmitter, the RSS-based database constructed by a specific radio transmitter is not suitable for different URTs. Since the test devices are usually different from the training devices, Hossain et al. [26] proposed a robust fingerprint method with received signal strength difference (RSSD) to eliminate the impact from hardware diversity. Dealing with the navigation system error causing by inertial sensors noises and biases, an approach applying wavelet packet de-noising to eliminate noises of the Micro-Electro-Mechanical Systems (MEMS) grade inertial sensors to effectively improve the positioning accuracy and ensure accurate and continuous positioning in the intelligent road service [27]. This paper dose research on the detection of URT with the fingerprint positioning technique. The RSSD information from two different APs is utilized as the characteristic parameter to establish the fingerprint database to achieve detection of the URT. Furthermore, RSSD-based fingerprint database is suitable for different URTs. In order to obtain more accurate sampling data stored in the fingerprint database, we also develop a new RSSD-based filtering method. The proposed filtering method and RSSD-VRKNN positioning algorithm are described in the following section.

1.2 Contribution and outline

To solve the issues mentioned above, our work mainly improves the parts of fingerprint database and positioning algorithm in the localization system to realize more accurate detection of the URT. Different from the previous works, our positioning scheme enhances the positioning accuracy by combining the virtual reference point and improved filtering method. The contributions of this paper are the following:

1) A new filtering method utilizing the RSSD information was presented to construct the fingerprint database. Moreover, the proposed RSSD-based filtering method improves the accuracy of fingerprint database, and it provides more accurate original data for positioning algorithm.

2) On the basis of the established RSSD-based database, we combined the KNN positioning algorithm with virtual reference point to propose a new RSSD-VRKNN positioning algorithm.

3) The feasibility of the proposed algorithm has been demonstrated by utilizing different emission strength and frequency between off-line database and on-line test target.

During the off-line phase, we first need to establish a RSSD-based fingerprint database with two RSS vectors from different APs. During the on-line phase, we apply KNN algorithm to obtain the virtual reference point and compare the virtual reference point with the reference points in the original database to find the new K reference points employed for the location estimation. Finally, we test the proposed RSSD-VRKNN algorithm in real test environment, where realistic measurements are performed. Experimental results certify that the proposed RSSD-VRKNN algorithm can not only detect the various unknown radio transmitters but also improve the positioning accuracy effectively compared with the existing work.

The rest of this paper is described as follows. We introduce the proposed positioning system for URT and the principle of KNN algorithm in Section 2. Next, the RSSD-VRKNN positioning technique is proposed in Section 3. In Sections 4 and 5, the positioning performance of RSSD-VRKNN is compared with RSSD-KNN and RSSD-WKNN algorithm through the simulation and experiments.

2 Positioning principle and algorithm

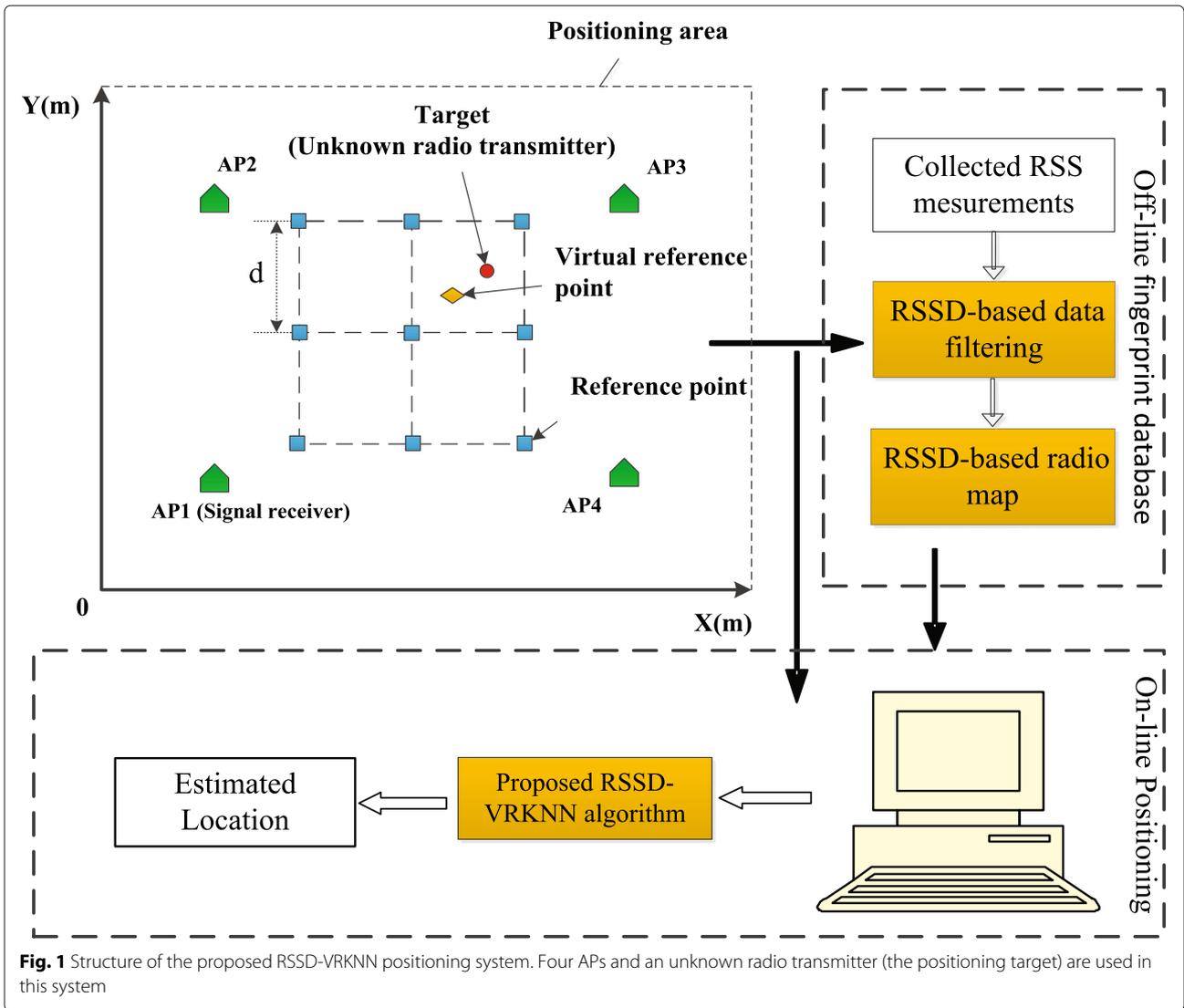
2.1 Proposed positioning system for URT

This sub-section introduces the proposed RSSD-VRKNN positioning system as described in Fig. 1. Our research is to solve the accurate detection of an URT. Then the signal receiver placed in the positioning area is used as the access point. On the basis of the conventional fingerprint positioning system, RSSD-based data filtering, RSSD-based fingerprint database, and RSSD-VRKNN algorithm are added as the new parts to realize the precise localization of the URT.

Four APs are applied as an example in our work, and d is defined as the grid distance between two reference points as shown in Fig. 1. Owing to the reflected performance of the radio signal, we usually do not choose the reference point near the wall. At the beginning of the RSSD-based database establishment, the APs collect RSS measurement from a known radio transmitter which is placed at each pre-planned reference point. The database mainly contains RSSD information and the location information of each reference point. $d_{i,j}$ is the distance between i -th reference point and j -th AP and it can be calculated by:

$$d_{i,j} = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}, \quad (1)$$

where i ($i=1, 2, \dots, n$) and j ($j=1, 2, \dots, l$) represent the index of reference point and AP, respectively. (X_i, Y_i) and



(X_j, Y_j) is the location coordinates of i -th reference point and j -th AP, respectively. In this paper, n is the number of reference point and it can be calculated by:

$$n = \left(\frac{L}{d} - 1\right) \cdot \left(\frac{H}{d} - 1\right), \tag{2}$$

where L and H are the length and width of the positioning area, respectively.

In off-line database establishment phase, RSSD-based data filtering process illustrated with yellow color is utilized to generate more precise RSSD data stored in fingerprint database. The role of the RSSD-based data filtering process is similar to that of the conventional system, but the detailed process is improved to deal with the collected RSS measurements and it will be expressed in next section. Moreover, RSSD-based database marked with yellow color has adaptability to

various unknown radio transmitters. If an URT entering the positioning area is different from the frequency and strength of the radio transmitter used for database establishment, the pre-established database will no longer be applicable. The RSSD-based database can solve this problem and greatly reduce the workload of establishing the database. In on-line positioning phase, the proposed RSSD-VRKNN algorithm is the core part of the proposed positioning system to improve the positioning accuracy. For this purpose, the conventional KNN algorithm is developed and improved in our work. The basic process of KNN algorithm is described in the next sub-section.

2.2 K-nearest neighbor algorithm

In this part, we introduce KNN algorithm which is one of the most famous algorithms in fingerprint positioning

technique. In this algorithm, Euclidean distance is used to represent the similarity between the real-time RSS and database, given by

$$Ed_i = \sqrt{\sum_{j=1}^l (\overline{\text{RSS}}_j - \overline{\text{RSS}}_{i,j})^2}, \quad (3)$$

where Ed_i represents the Euclidean distance between i -th reference point and the positioning target, $\overline{\text{RSS}}_j$ is the real-time average RSS value of j -th AP, and $\overline{\text{RSS}}_{i,j}$ is the average RSS of i -th reference point from the j -th AP. If the Euclidean distance between the positioning target and reference point in the database is shorter, the pattern between the two points is very similar, and vice versa. When the positioning target enters the positioning area, the real-time RSS is compared with the data of each reference point in the database, and the location of positioning target is estimated by selecting K reference points which Euclidean distances are minimum. After the process described above, the location of positioning target can be obtained by:

$$\begin{cases} x = \frac{\sum_{i=1}^K X_i}{K} \\ y = \frac{\sum_{i=1}^K Y_i}{K} \end{cases}, \quad (4)$$

where (x, y) is the estimated location of positioning target.

3 Proposed RSSD-VRKNN positioning technique

Unlike the previous work on the localization of signal receiver, our research is to place the signal receiver to find the URT. In this section, we introduce how the proposed positioning system is used to estimate the location of an URT. The basic process of the proposed RSSD-VRKNN positioning system is as shown in Fig. 2. The proposed RSSD-VRKNN fingerprint positioning system also consists of two parts, off-line database establishment phase and on-line positioning phase.

3.1 RSSD-based data filtering and radio map establishment

In this subsection, the process of the RSSD-based radio map establishment is described. Next, we first explain and derive the relationship between the RSSD and the influence of diverse radio transmitters. Moreover, we demonstrate why there is no influence of diverse radio transmitters when using the RSSD-based information.

According to the log-distance path loss model, we denote that $P(d_0)$, $P(d_p)$, and $P(d_q)$ are the RSS at distance d_0 , d_p , and d_q from the radio transmitter to p -th and q -th AP, respectively. d_0 is the reference distance. Thus, the relationship among d_0 , $P(d_0)$, $P(d_p)$, and $P(d_q)$ can be expressed as:

$$P(d_p) = P(d_0) - 10 \cdot \alpha \cdot \log\left(\frac{d_p}{d_0}\right) + \chi_p \quad (5)$$

and

$$P(d_q) = P(d_0) - 10 \cdot \alpha \cdot \log\left(\frac{d_q}{d_0}\right) + \chi_q, \quad (6)$$

where α is the pass loss component. χ_p and χ_q represent the shadowing effect, which obey zero-mean Gaussian distribution ($\chi \sim N(0, \sigma^2)$). The typical values for $d_0 = 1m$, $P(d_0) = 10dB$, $\sigma = 5.2dB$, and $\alpha = 1.8$ in [28]. The free space propagation model [29] can be explained as:

$$P(d_0) = 10 \log\left(\frac{P_t G_t G_j \lambda^2}{(4\pi)^2 d_0^2 L_j}\right) \quad (7)$$

where P_t is the power of radio transmitter, G_t is transmitted antenna gain, G_j is the j -th AP's antenna gain, L_j is the system loss factor, and λ is the transmitter carrier's wavelength.

According to (5), (6), and (7), we can obtain the RSS of p -th and q -th AP respectively as:

$$P(d_p) = 10 \log\left(\frac{P_t G_t G_p \lambda^2}{(4\pi)^2 d_0^2 L_p}\right) - 10 \cdot \alpha \cdot \log\left(\frac{d_p}{d_0}\right) + \chi_p \quad (8)$$

and

$$P(d_q) = 10 \log\left(\frac{P_t G_t G_q \lambda^2}{(4\pi)^2 d_0^2 L_q}\right) - 10 \cdot \alpha \cdot \log\left(\frac{d_q}{d_0}\right) + \chi_q. \quad (9)$$

According to (8) and (9), the RSSD between p -th and q -th AP can be calculated by:

$$P(d_p) - P(d_q) = 10 \log\left(\frac{G_p L_q}{G_q L_p}\right) + 10 \log\left(\frac{d_q}{d_p}\right) + (\chi_p - \chi_q). \quad (10)$$

As we can see clearly from (10), RSSD is not affected by the diversity of radio transmitters. Therefore, if remaining the APs same, we don't need to change and establish the fingerprint database again, which not only can greatly reduce the workload to establish fingerprint database but also can improve the positioning accuracy and system stability.

Due to the impacts of reflection, scattering, shadowing, and changing scenario in radio signal propagation, the sampling RSS will be not accurate. As shown in Fig. 3, it reveals the RSSD distribution of AP12 with no data filtering. Next, we introduce how the proposed RSSD-based data filtering is utilized to construct the new fingerprint database. As mentioned in the previous section, the sampling RSS vector of i -th reference point from j -th AP is

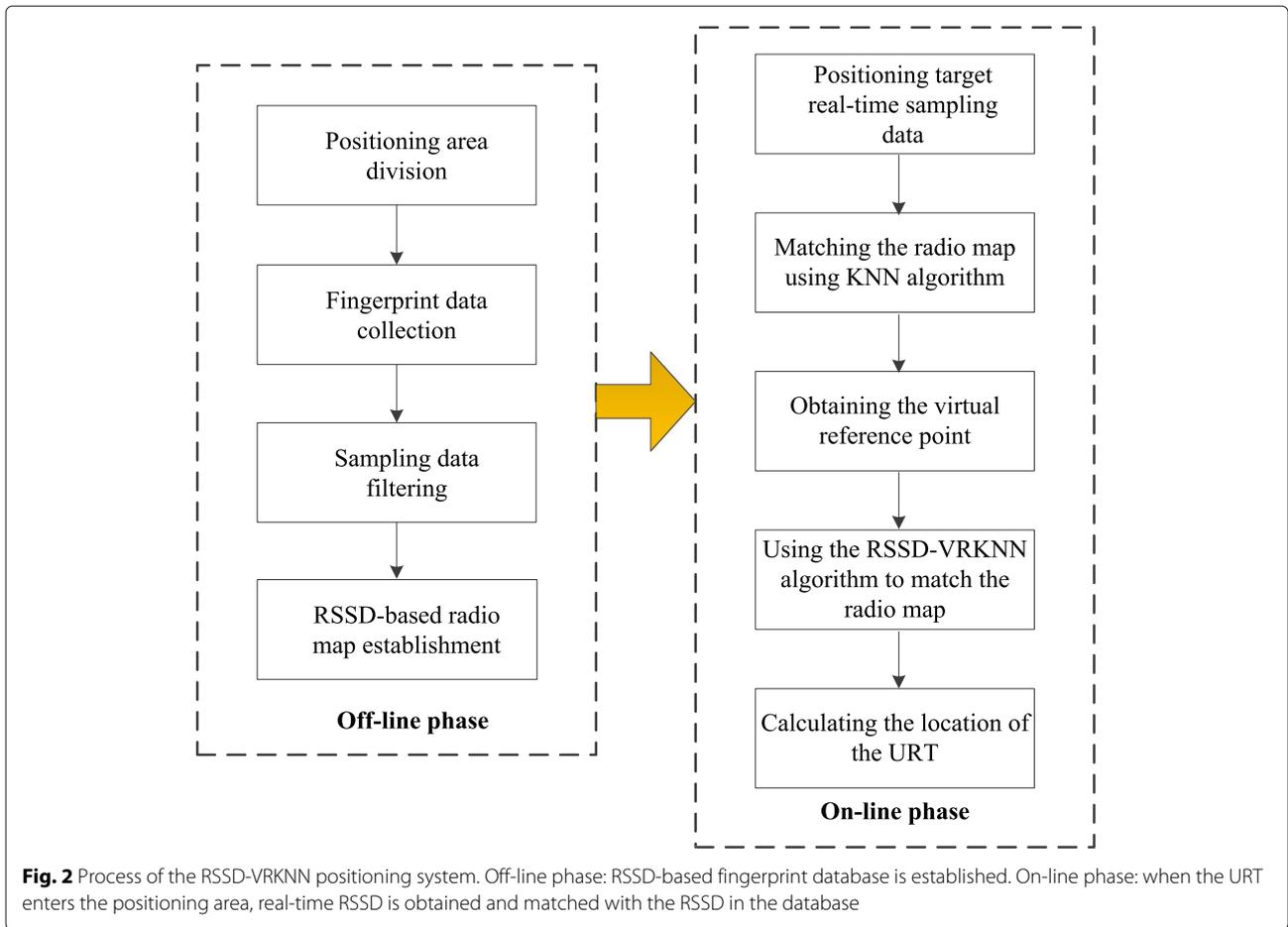


Fig. 2 Process of the RSSD-VRKNN positioning system. Off-line phase: RSSD-based fingerprint database is established. On-line phase: when the URT enters the positioning area, real-time RSSD is obtained and matched with the RSSD in the database

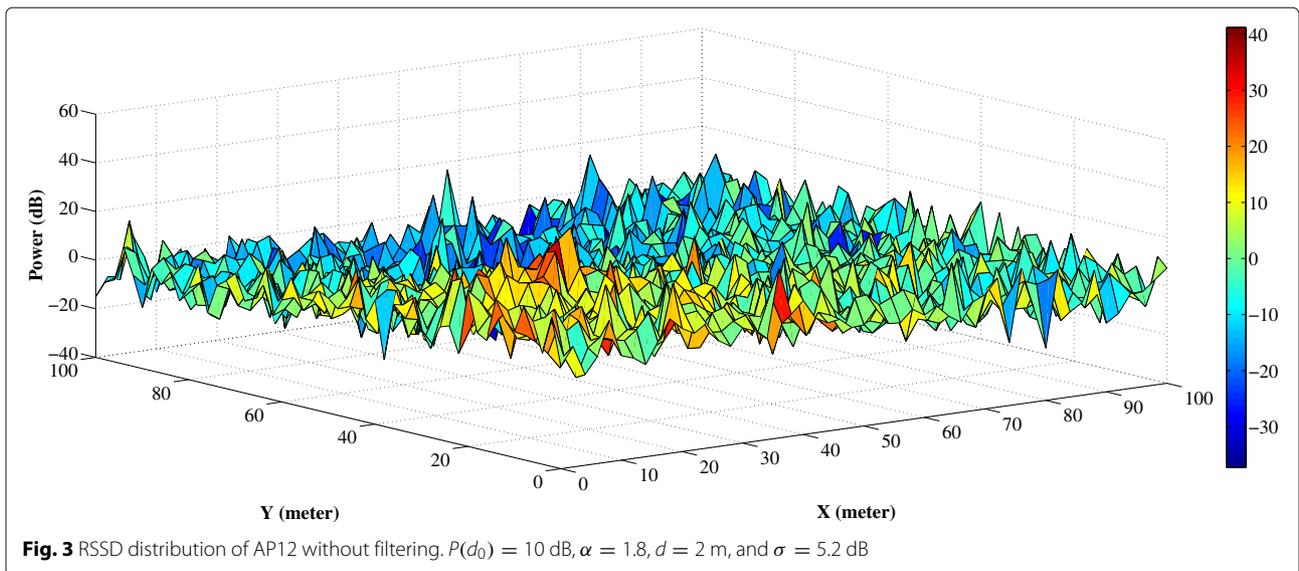


Fig. 3 RSSD distribution of AP12 without filtering. $P(d_0) = 10$ dB, $\alpha = 1.8$, $d = 2$ m, and $\sigma = 5.2$ dB

$AVR_{i,j}$. In the same way, the sampling RSSD vector of two different APs can be expressed by:

$$AVR_{i,p} = \left(RSS_{i,p}^1, RSS_{i,p}^2, \dots, RSS_{i,p}^m \right) \quad (11)$$

and

$$AVR_{i,q} = \left(RSS_{i,q}^1, RSS_{i,q}^2, \dots, RSS_{i,q}^m \right). \quad (12)$$

From (11) and (12), the sampling RSSD vector of i -th reference point between p -th AP and q -th AP can be calculated by:

$$RSSD_{i,pq} = \left[\left(RSS_{i,p}^1 - RSS_{i,q}^1 \right), \left(RSS_{i,p}^2 - RSS_{i,q}^2 \right), \dots, \left(RSS_{i,p}^m - RSS_{i,q}^m \right) \right], \quad (13)$$

where p and q are the index of AP ($p, q = 1, 2, \dots, l$). The combination of p and q is "12, 23, 34, 41" in this paper. Then, the average of RSSD within a period time can be obtained by:

$$\overline{RSSD}_{i,pq} = \frac{\sum_{m=1}^t RSSD_{i,pq}^m}{t}, \quad (14)$$

where t is the number of sampling data. When obtaining the average of $RSSD_{i,pq}$, we calculate the Euclidean distance between each sampling $RSSD_{i,pq}^m$ and the average value $\overline{RSSD}_{i,pq}$. After that, an appropriate threshold is set to compare with each Euclidean distance of sampling $RSSD_{i,pq}^m$. If the threshold is exceeded, the sampling $RSSD_{i,pq}^m$ is deleted. Then, the new average will be calculated again until all the sampling $RSSD_{i,pq}^m$ data meet the requirement. According to the measure experience, the threshold is set with 5 dB in this paper. Finally, the calculated average RSSD value is stored in the database. The flowchart of RSSD-based fingerprint database establishment is shown in Fig. 4. The yellow shaded part indicates the proposed data filtering process. The structure of the improved database based on RSSD is shown as Fig. 5. The filtered RSSD illustrated with yellow shaded part composes the fingerprint database, and the improved database has higher reliability and accuracy compared with conventional fingerprint database. Taking AP12 as an example, Fig. 6 shows the RSSD distribution with proposed RSSD-based filtering. To reflect the influence of the proposed filtering method on the sample accuracy of fingerprint database, we use the proposed RSSD-based data filtering as expressed from (11)–(14) to make a comparison with traditional mean filtering method as shown in Fig. 7. In Fig. 7, x -coordinate represents 10 different location serial numbers randomly selected from the positioning area. At each location point, 50 RSSD samples collected by different AP combinations were recorded

respectively and constituted samples. The filtering performance was evaluated with the sample standard deviation expressed by:

$$\sigma_s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (RSSD_{f,i} - \overline{RSSD})^2}, \quad (15)$$

where \overline{RSSD} is mean value of the samples and $RSSD_{f,i}$ is the filtering value of i -th AP combination. The results show that the proposed filtering method is more accurate and has stronger anti-interference.

3.2 Proposed RSSD-VRKNN fingerprint positioning algorithm

In fingerprint positioning technique, we usually expect to get the location coordinates and RSS information of the surrounding points, which are around the positioning target. With this idea, an improved KNN algorithm based on RSSD and virtual reference point is proposed in this paper. We assume that (x, y) and (x_j, y_j) are the the location coordinate of the URT and AP, respectively. First, the RSSD-based fingerprint database needs to be established. According to (3) and (14), the Euclidean distance of RSSD can be obtained by:

$$SSDEd_i = \sqrt{\sum (\overline{RSSD}'_{pq} - \overline{RSSD}_{i,pq})^2}, \quad (16)$$

where \overline{RSSD}'_{pq} is the real-time average RSSD value of two different APs. Second, we use KNN algorithm to obtain the initial location of the positioning target, which is expressed by:

$$\begin{cases} RSSDX_i = \frac{\sum_{i=1}^k X_i}{K} \\ RSSDY_i = \frac{\sum_{i=1}^k Y_i}{K} \end{cases}. \quad (17)$$

In the proposed RSSD-VRKNN algorithm, we define the initial location of positioning target as the virtual reference point. As with other reference points in the database, we also need to know RSS and location coordinates of the calculated virtual reference point. According to the log-normal shadowing model, RSS is mainly related to the distance. The closer to the virtual reference point, the closer the reference point is to the scene information of the virtual point. Therefore, we apply the Euclidean algorithm of distance correlation to select K -nearest neighbors of the reference points to calculate the RSS of virtual reference point in the database. The Euclidean algorithm of distance

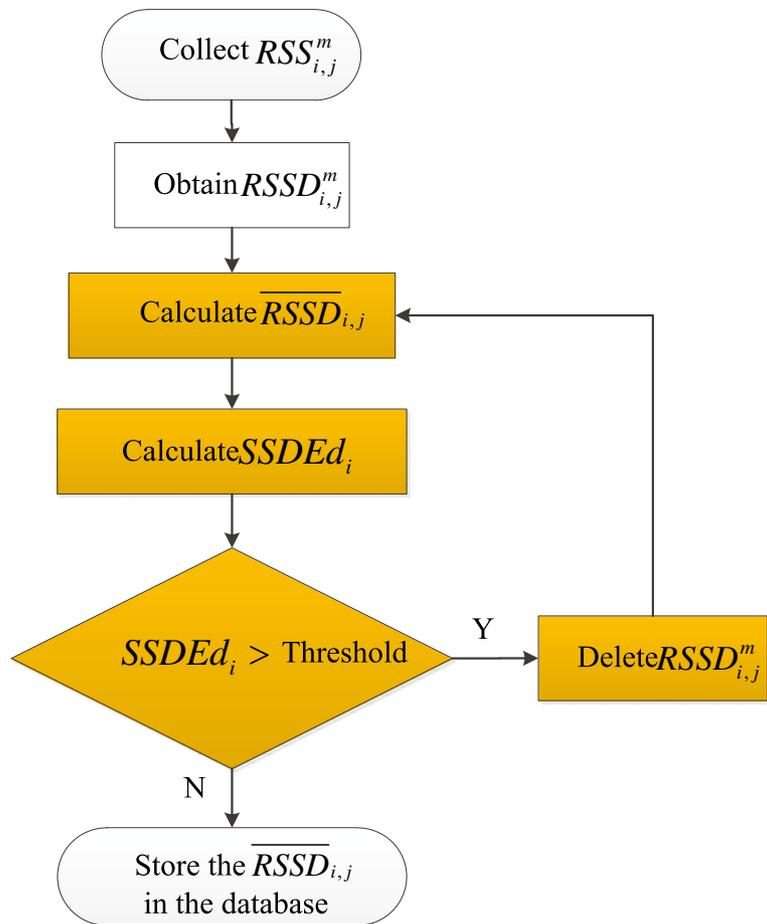


Fig. 4 Flowchart of proposed RSSD-based fingerprint database establishment. Take the sampling RSS of i -th reference point from j -th AP for example

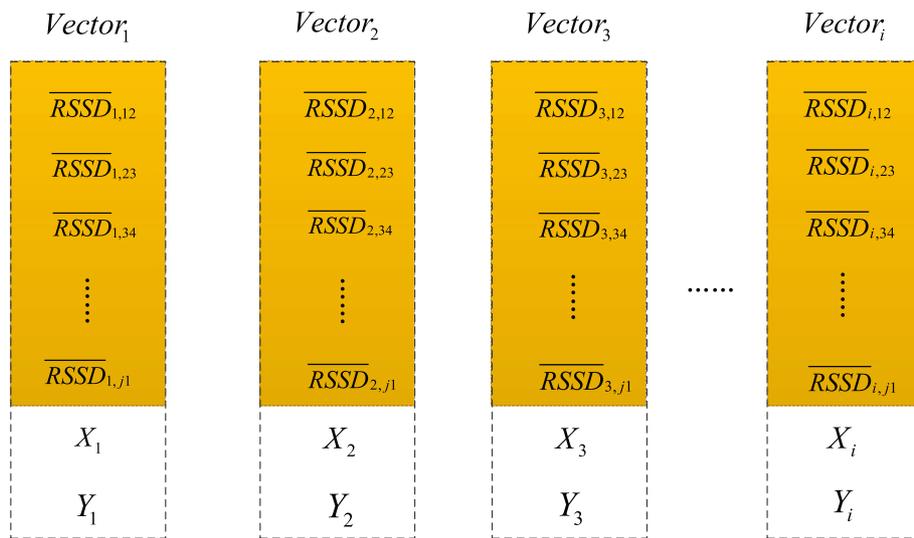
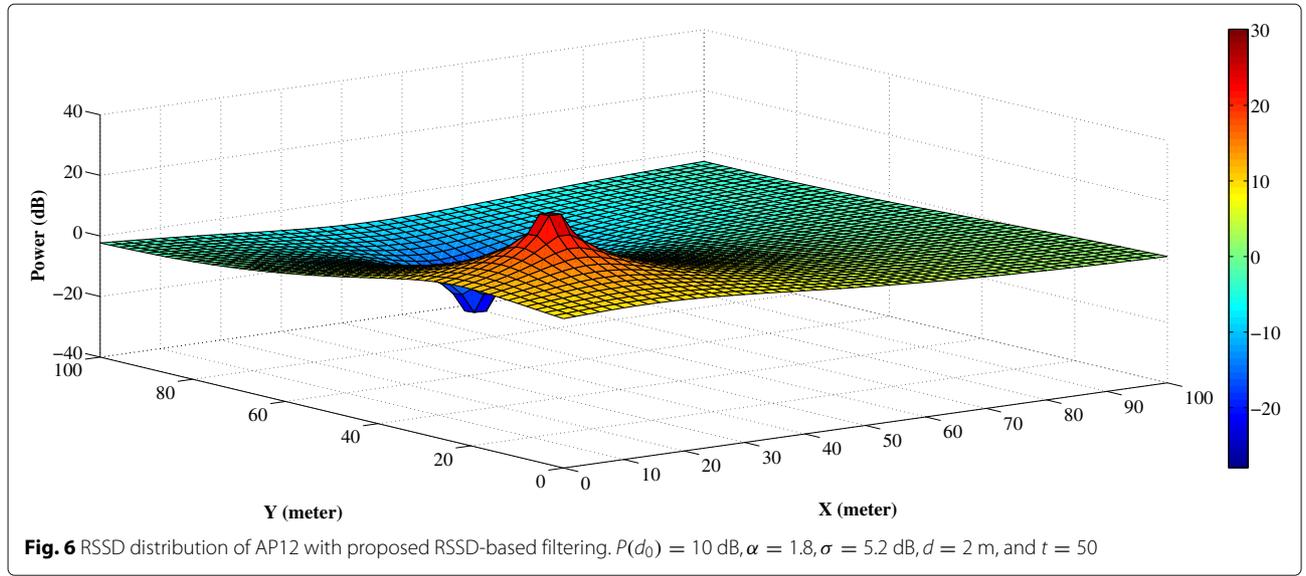


Fig. 5 Structure of the proposed RSSD-based database. $vector_i$ represents the data vector of i -th reference point, which contains coordinates and RSSD information



correlation can be expressed as:

$$DEd_i = \sqrt{\sum_{j=1}^l (d_{v,j} - d_{i,j})^2}, \tag{18}$$

where $d_{v,j}$ is the distance between the virtual reference point and j -th AP. In the same way as the RSS-based nearest neighbor algorithm, we select K reference points of the smallest DEd to calculate the RSS of virtual reference point. Since RSS is related to distance, we use distance-based weight algorithm to estimate the average RSS value

of virtual reference points and it can be expressed by:

$$\overline{RSS}_{v,j} = d_{v,j} \cdot \left(\frac{\sum_{i=1}^k d_{i,j} \cdot \overline{RSS}_{i,j}}{\sum_{i=1}^k d_{i,j}} \right), \tag{19}$$

where $d_{v,j}$ is the distance between the virtual reference point and j -th AP and it can be obtained from (1) and (17). $\overline{RSS}_{i,j}$ is the average RSS value between i -th reference point and j -th AP after filtering. When completing

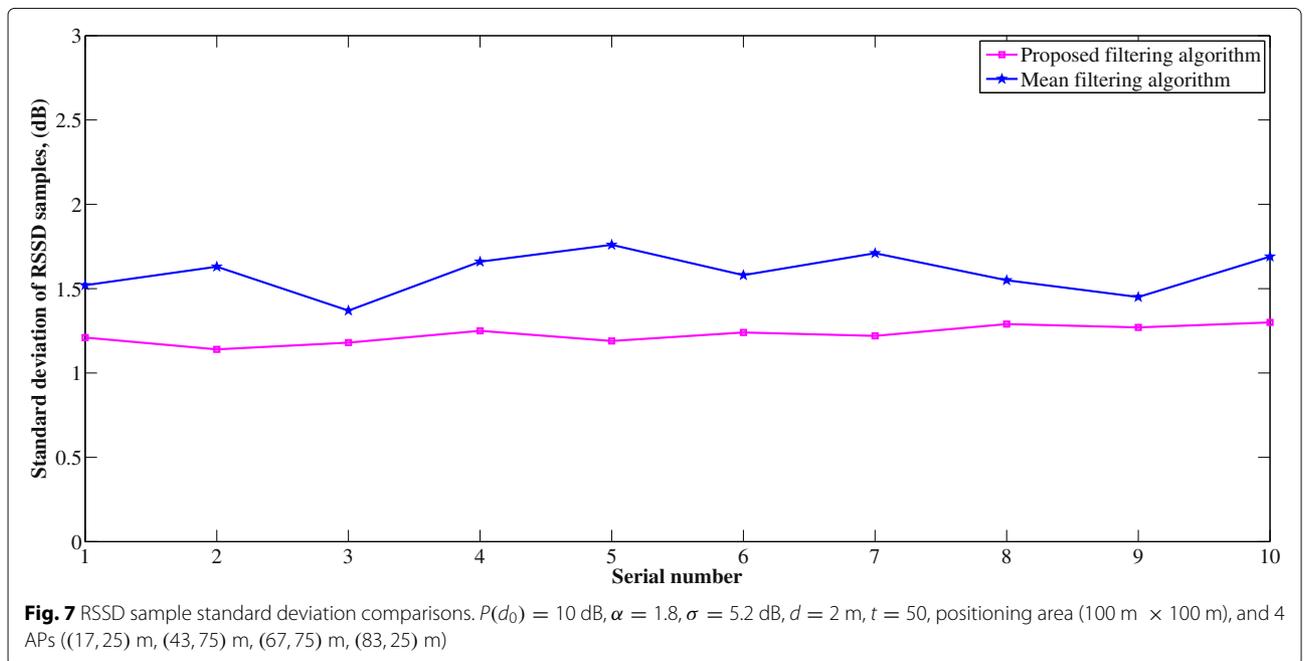


Table 1 Simulation parameters

Parameter	Symbol	Value
Positioning area dimension	$L \times H$	100 m \times 100 m
Grid distance	d	2 m
Pass loss component	α	1.8
Reference distance	d_0	1 m
RSS at distance d_0	$P(d_0)$	10 dB, 30 dB
No. of APs	l	4
AP position	(x_j, y_j)	AP1 (17, 25) m AP2 (43, 75) m AP3 (67, 75) m AP4 (83, 25) m
Variance of RSS	σ^2	5.2 ² dB
Number of sampling data	t	50

the above process, we obtain the RSS and location information of the virtual reference point. Finally, the obtained RSSD of the virtual reference point into the database is added to generate a new database. Then, we utilize the RSSD-based KNN algorithm to select the new K reference points to calculate the location of positioning target.

4 Simulation analysis

4.1 Simulation setup

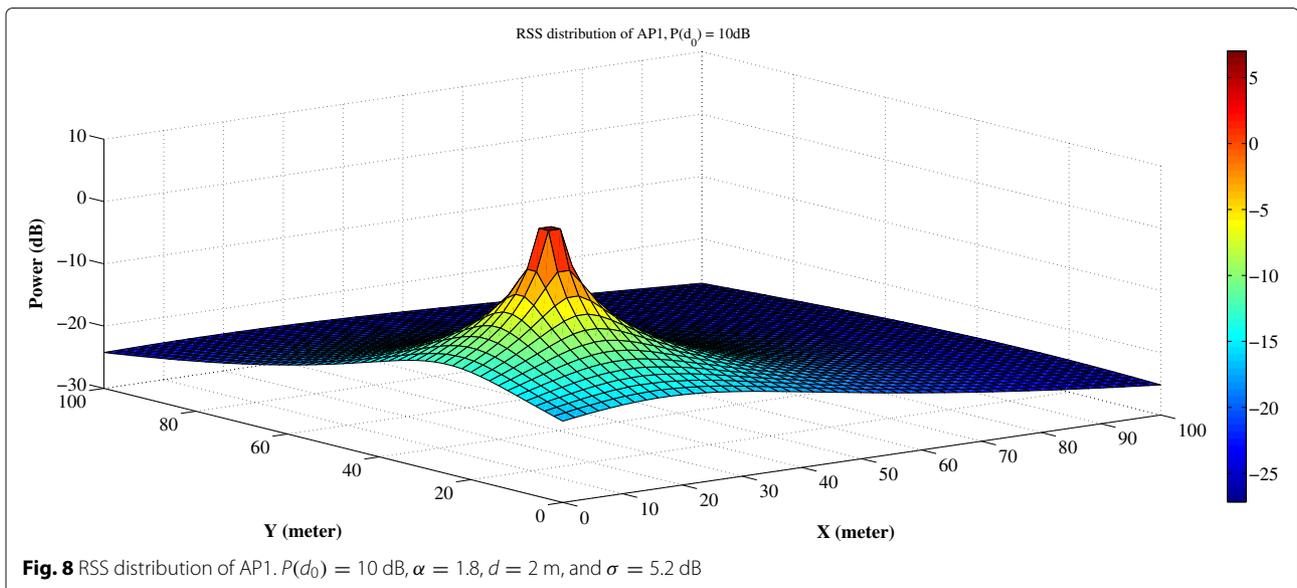
Different positioning algorithms are numerically simulated in MATLAB 2014a software environment to compare with the proposed RSSD-VRKNN algorithm. The simulation RSS measurements used for the off-line RSSD database phase and on-line positioning target need to be generated. To imitate the real positioning scenario, we

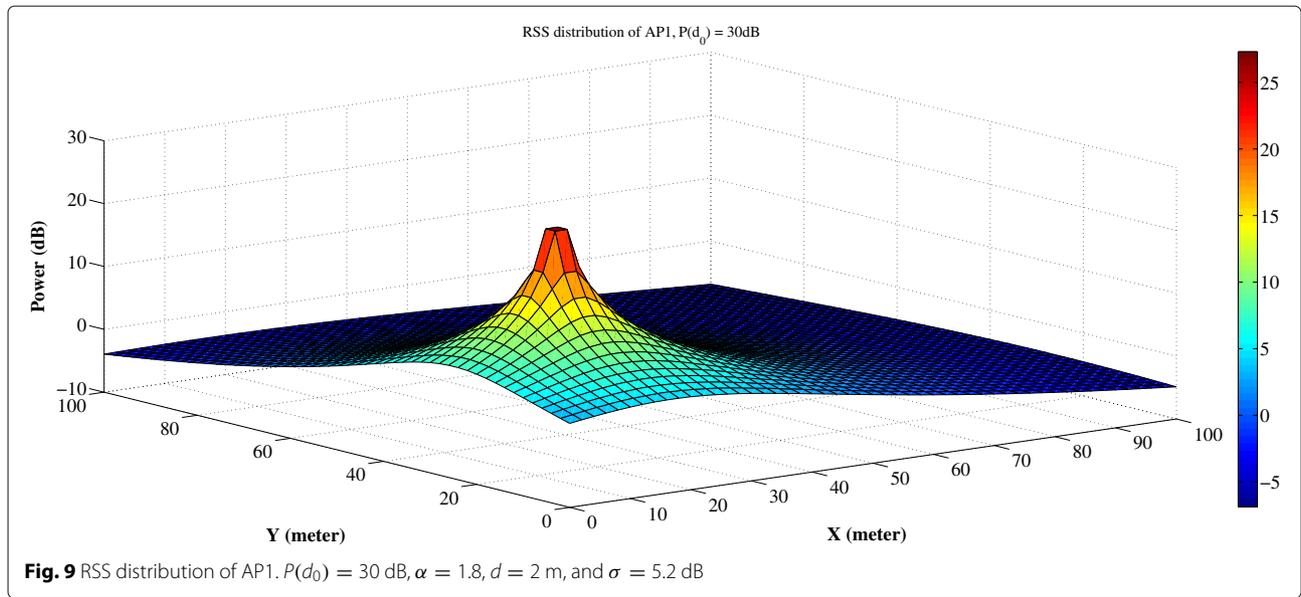
chose a well-known log-normal shadowing model [28] to generate the random RSS measurements. The random variable $P(d_{i,j})$ represented the $RSS_{i,j}$ measurement can be supposed as Gaussian with $P(d_{i,j}) \sim N(\overline{P(d_{i,j})}, \sigma^2)$, where $\overline{P(d_{i,j})} = P(d_0) - 10 \cdot \alpha \cdot \log(d_{i,j}/d_0)$ is the mean RSS measurement value, σ^2 is the variance of the shadowing. The typical simulation parameters [28] are described as in Table 1.

4.2 Comparison of RSSD-based and RSS-based

As described in (7), RSS-based fingerprint database depends on the power and frequency of the radio transmitter. In the off-line database establishment phase, the database is different with the diversity of radio transmitters. Unlike the previous fingerprint positioning technique, the positioning target is an unknown radio transmitter in this paper, which frequency and strength are all unknown. Therefore, the RSS-based database is not suitable for positioning the URT. From (10), we can obtain that the RSSD-based database is not affected by the strength and frequency of the radio transmitter. It means that we only need to establish one off-line database to be able to adapt to the localization of different unknown radio transmitters.

As shown in Figs. 8 and 9, RSS distribution of different radio transmitters are significantly different. Therefore, the RSS-based off-line database is not suitable for detection of the URT. However, as can be seen from Figs. 10 and 11, the RSSD-based database is not affected by the comparison of the RSS-based database, when the radio transmitter is different. Therefore, this paper utilize RSSD-based fingerprint database to realize the localization of an URT.



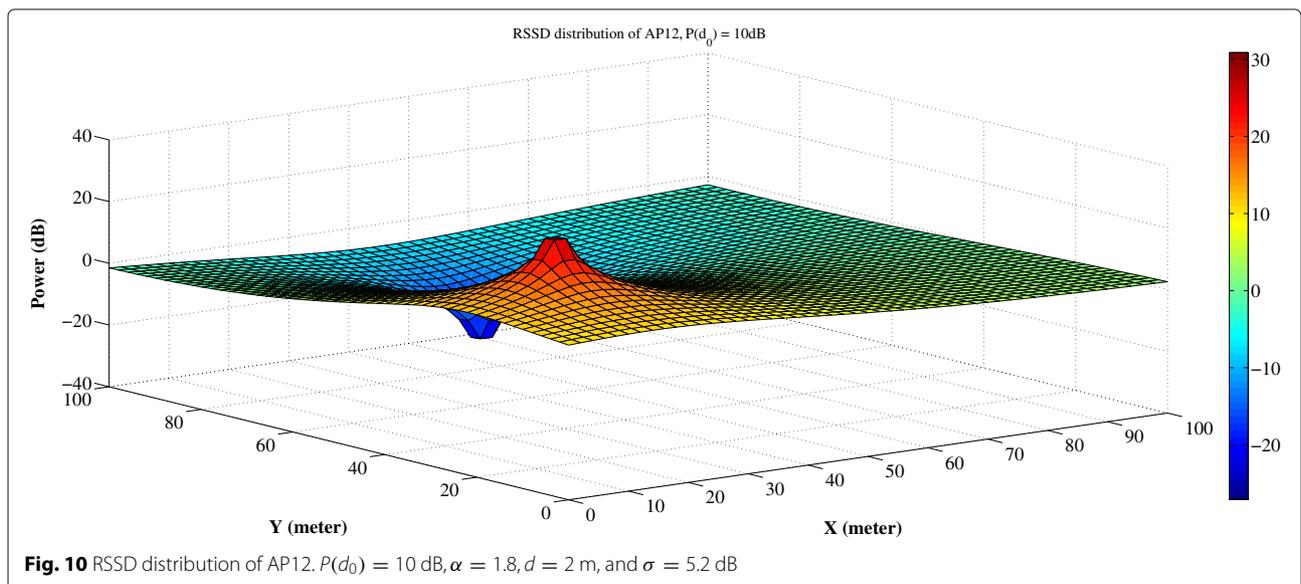


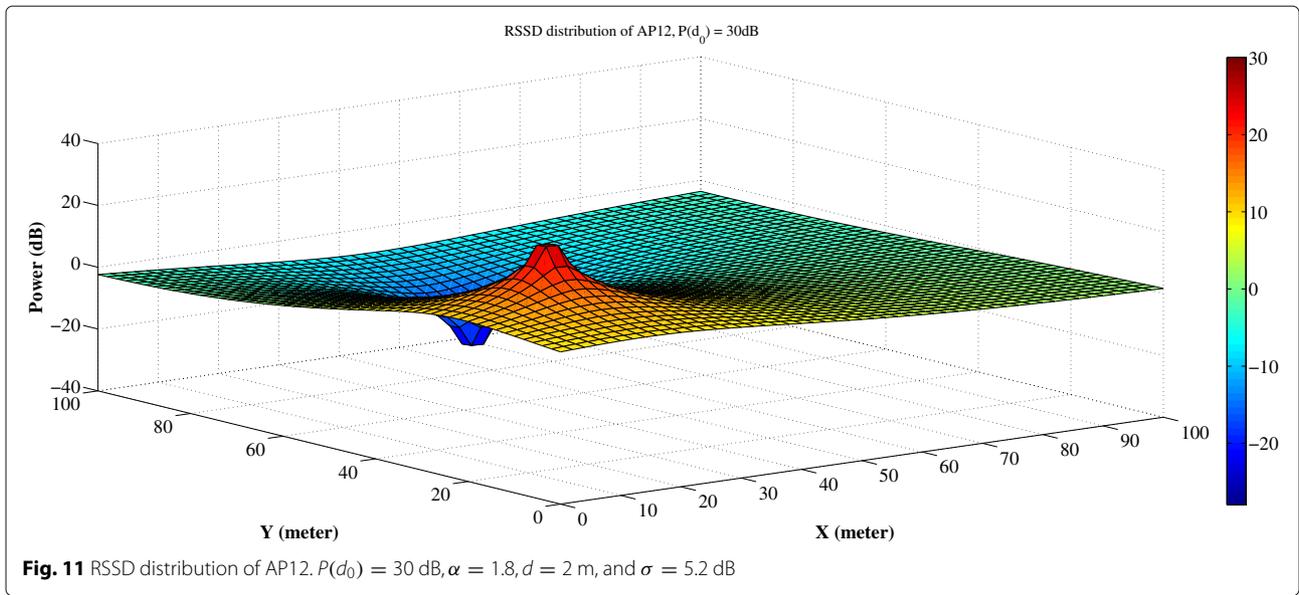
4.3 Performance of proposed RSSD-VRKNN positioning error

The performance of the proposed RSSD-VRKNN algorithm was testified by simulation as shown in Fig. 12, which consists of 100 single-target locations randomly chosen from the positioning area of $L(100\text{ m}) \times H(100\text{ m})$, 100 estimated target locations and 4 AP locations. It should be noted that we consider only one target positioning problem in this paper. The detection of multiple unknown radio transmitters is left for our future work. As shown in Fig. 12, four APs were used in total and identified with the “ Δ ” mark. The locations of the APs in the simulation are fixed at (17, 25) m,

(43, 75) m, (67, 75) m, and (83, 25) m, respectively. As described above, the positioning targets are randomly selected to evaluate the performance of the proposed algorithm.

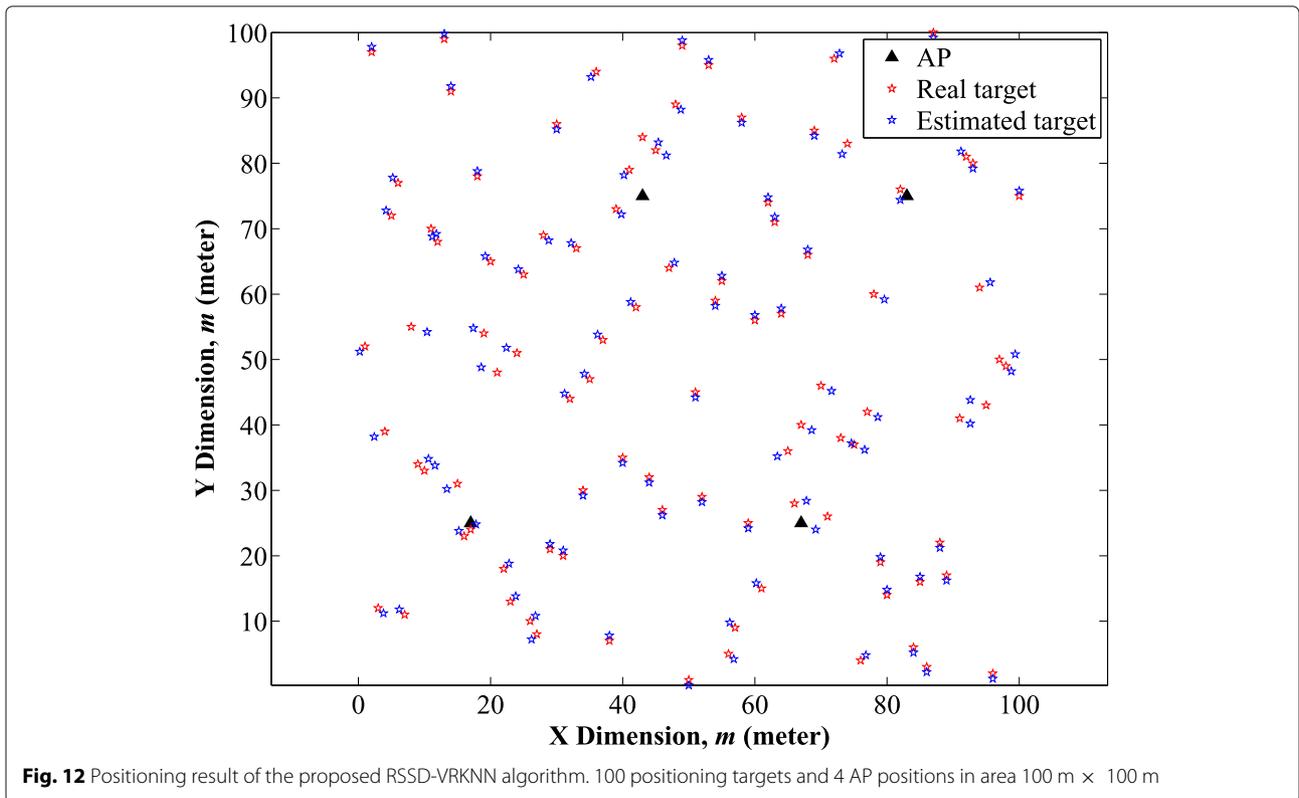
In order to illustrate the superiority of the proposed RSSD-VRKNN algorithm, KNN algorithm, and WKNN algorithm are selected to be compared. Figure 13 shows the root mean square error (RMSE) of the three algorithms, when the number of positioning targets changes from 10 to 100. As we can see from the simulation results, the proposed RSSD-VRKNN algorithm has higher positioning accuracy than the other two algorithms. The RMSE of the proposed RSSD-VRKNN algorithm is stable

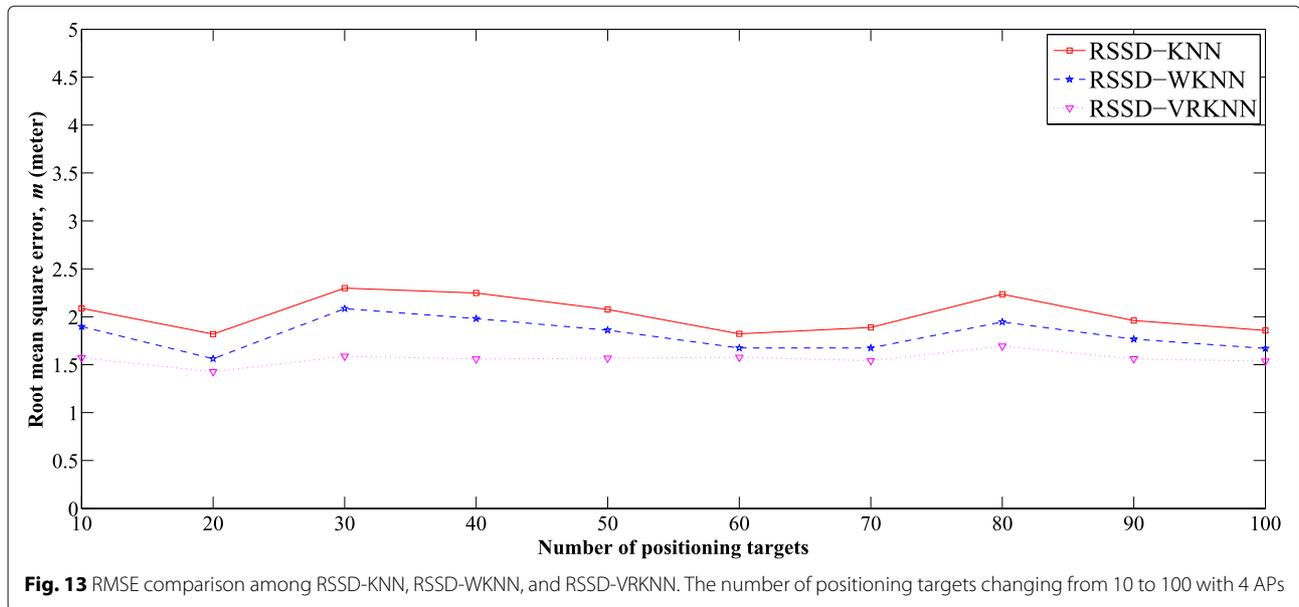




at about 1.5m . Taking 20, 50, and 100 positioning targets as examples, Table 2 shows the performance comparison of different algorithms with location error. Due to uncertain factors such as the randomly selected positioning targets, changing scenario and AP location, RMSE of different numbers of positioning targets has some volatility. However, the proposed algorithm has better stability than

the other two algorithms. As shown in Fig. 14, there is an obvious difference in the tendency of the RMSE among the RSSD-KNN, RSSD-WKNN, and RSSD-VRKNN when the number of AP changes from 3 to 5. Comparing with the conventional RSSD-KNN and RSSD-WKNN techniques, the RMSE with three APs shows better performance than proposed RSSD-VRKNN. This is because few





APs lead to a lack of information in the RSS of the reference point. When the number of APs increases to 4, the proposed algorithm attains better localization performance than RSSD-KNN and RSSD-WKNN. However, the RMSE with 4 APs and 5 APs is almost the same. The comparison of location errors with different APs is revealed in Table 3. This indicates that the positioning accuracy will not be enhanced when the number of APs increases consistently. Therefore, taking the hardware consumption and cost into consideration, this paper uses 4 APs to meet the positioning requirements.

5 Experimental results

5.1 Experimental environment and conditions

In order to further verify the reasonability of the proposed algorithm, fingerprint database was established and some test positioning targets were selected in the actual environment. The experiment was conducted to evaluate its performance in a real office and corridor, which is located on the first floor of National Radio Monitoring Center, Beijing. The office and corridor have a dimension of $14.4 \text{ m} \times 8.4 \text{ m}$ and $14.4 \text{ m} \times 2.1 \text{ m}$, respectively, and the total area is 151.2 m^2 as shown in Fig. 15. In addition, the office has brick structure with four windows and two

doors in the wall. There is no partition or compartment in the office and it mainly consists of six rows of desks, chairs, and computers. During the test, the staff are free to enter and leave frequently.

According to the instruction of simulation result, we deployed four SA44B (Signal Hound Co. Ltd.) measuring receivers as APs in the office, and the location coordinates are $(3, 2)m$, $(8, 5)m$, $(3, 9)m$, and $(8, 12)m$, respectively. In the off-line database establishment phase, the grid distance was selected with 1.8 m. The radio transmitter TFG6300 (SUING Co. Ltd.) used in the experiment is adjustable in frequency and strength. Here, we define “frequency/strength” as the frequency and strength of the radio transmitter. In the off-line database establishment phase, the frequency and strength of the radio transmitter is “300 MHz/20 dB” and “1 GHz/20 dB” respectively. In the on-line positioning phase, we selected the radio transmitter with the frequency of 300 Mhz and the strength of 20 dB and 13 dB, respectively as the test positioning target. In the positioning area, 50 test points with 1.5-m grid distance were selected for localization test.

5.2 Performance comparisons

Next, we compared the performance of RSSD-VRKNN with RSSD-KNN and RSSD-WKNN. In our test, the number of K was set to 4 and the number of sampling RSS was set to 50. The mean location error and cumulative distribution function (CDF) of location error were utilized as the key indicators for performance evaluation. The mean location error represents the average value of all the positioning target that deviated from the real position. The CDF is characterized by the distribution of the location errors.

Table 2 The probability of location error ($\leq 1.5 \text{ m}$) with different numbers of positioning targets

Number of positioning targets	RSSD-KNN	RSSD-WKNN	RSSD-VRKNN
20	45%	60%	80%
50	44%	56%	72%
100	47%	59%	75%

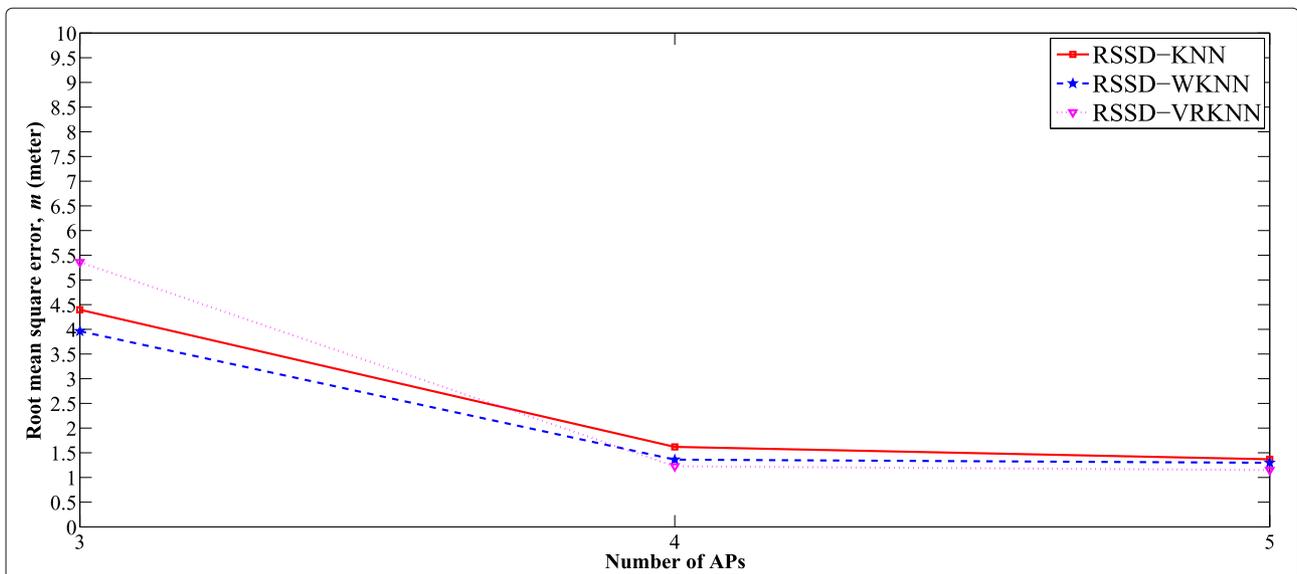


Fig. 14 RMSE comparison among RSSD-KNN, RSSD-WKNN, and RSSD-VRKNN. The number of APs varies from 3 to 5 with 100 positioning targets

The mean location errors of RSSD-KNN, RSSD-WKNN, and RSSD-VRKNN algorithms are shown as in Table 4. The grid distance of the positioning area is 1.8 m. In order to represent the performance of the proposed algorithm, two databases with different frequencies and the same strength were established, which are 300 Mhz/20 dB and 1 Ghz/20 dB. The combination of frequency and strength of the test positioning targets were 300 Mhz/20 dB and 300 Mhz/13 dB, respectively. According to the frequency and strength parameters of reference points and test positioning targets, the following four different experiment results can be obtained:

1) The same frequency and strength: the mean location error by RSSD-VRKNN is 1.12 m. In comparison, the mean location error by RSSD-KNN and that by RSSD-WKNN are 1.43 m and 1.35 m, respectively.

2) The same frequency and different strength: the mean location errors of RSSD-KNN, RSSD-WKNN, and RSSD-VRKNN are 1.52 m, 1.38 m, and 1.17 m.

3) The different frequency and same strength: compared with RSSD-KNN and RSSD-WKNN, the mean location error of proposed RSSD-VRKNN algorithm is 1.21 m. The

mean location errors of RSSD-KNN and RSSD-WKNN are 1.48 m and 1.32 m, respectively.

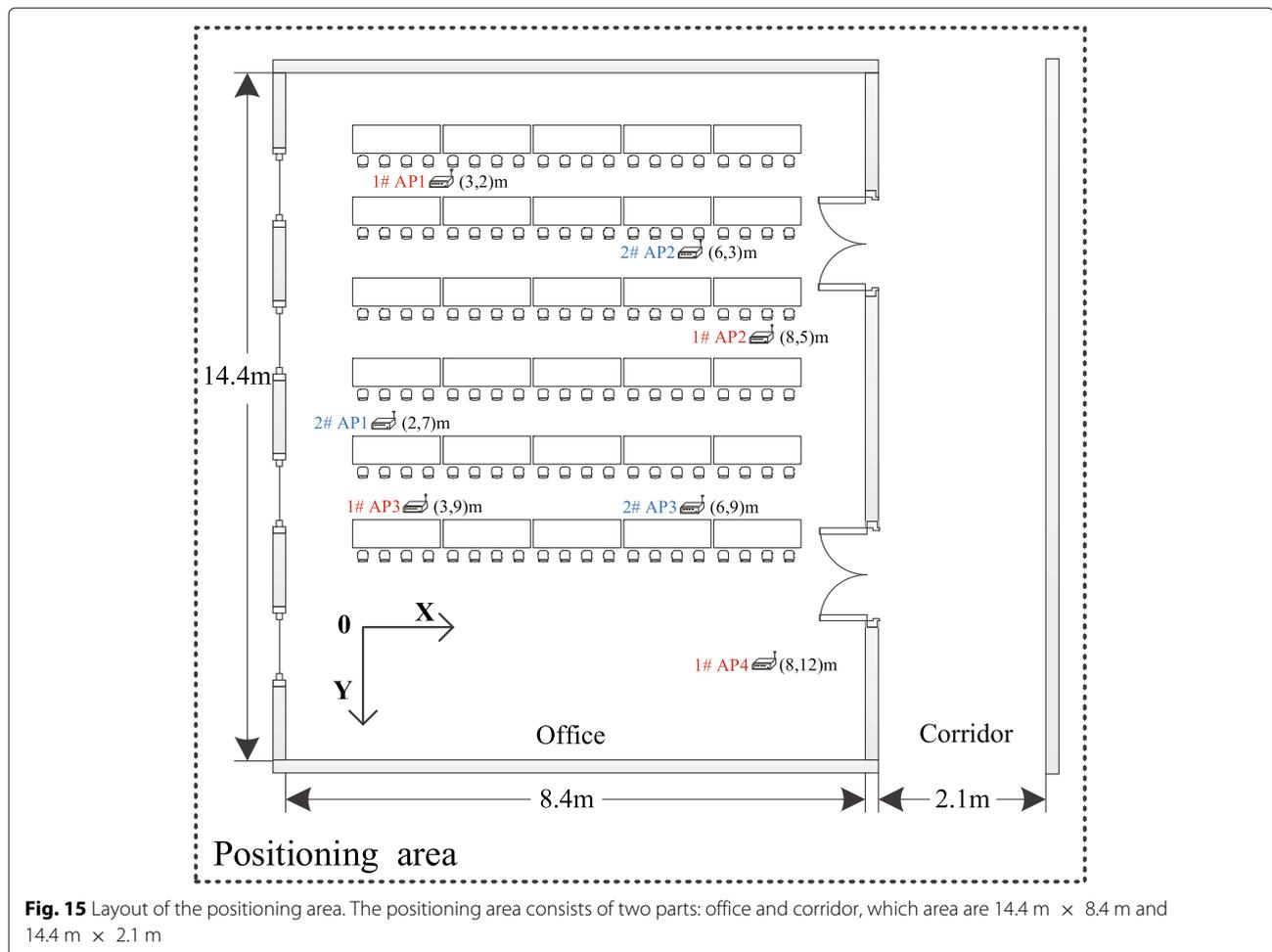
4) The different frequency and different strength: the location mean location errors of RSSD-KNN and RSSD-WKNN are 1.45 m and 1.33 m, respectively. The mean location error by RSSD-VRKNN is 1.15 m.

In summary, the positioning performance of proposed RSSD-VRKNN algorithm is better than RSSD-KNN and RSSD-WKNN from the above experimental results. Due to the impact of different positioning environment such as the movement of human being, the positioning accuracy is affected in each experiment. However, it is obviously obtained that the RSSD-based fingerprint database is not affected by the frequency and strength of the unknown radio transmitter. Therefore, the proposed RSSD-VRKNN can achieve accurate detection of the unknown radio transmitter.

Due to the RSSD-based database has been demonstrated to be suitable for different frequency and strength of the radio transmitter. Hence, we just selected the 1 GHz/20 dB database and 300 MHz/13 dB test positioning target to compare the CDF of RSSD-KNN, RSSD-WKNN, and RSSD-VRKNN as shown in Fig. 16. The other circumstances are similar. The number of the test positioning targets is 50 with 1.5-m grid distance. The curves of RSSD-KNN and RSSD-WKNN are very close. Considering the 80th percentile, RSSD-VRKNN has a location error of under 2 m. In comparison, the location errors of RSSD-KNN and RSSD-WKNN are 3.4 m and 3 m, respectively. Obviously, the proposed RSSD-VRKNN has the better performance than the other two algorithms in terms of CDF.

Table 3 The probability of location error (≤ 1.5 m) with different numbers of APs

Number of APs	RSSD-KNN	RSSD-WKNN	RSSD-VRKNN
3	36%	43%	28%
4	42%	56%	72%
5	48%	61%	77%



6 Conclusions

To achieve the detection of an unknown radio transmitter and improve the accuracy, this paper has proposed a RSSD-VRKNN algorithm that combines the RSSD parameter and virtual reference point. The RSSD-based database is utilized to adjust the diversity of radio transmitters. In this way, we only need to establish an off-line fingerprint database, which saves a lot of work due to the establishment of different databases to adapt to different radio transmitters. With the proposed RSSD data-filtering method, the RSSD measurements stored in the fingerprint database is more accurate. An enhanced KNN algorithm developed with adding virtual reference point improved

the detection accuracy for the URT. In addition, a large number of computer simulations were conducted to prove the correctness and validity of the proposed algorithm in different frequency and strength between the radio transmitter and database. The experiment results show that the mean location error of the proposed novel can attain the positioning accuracy of 1.12~1.21 m, which is about 15% and 20% lower than that of RSSD-WKNN and RSSD-KNN, respectively. The influence of AP's layout and detection of multiple URTs will be left to our future work. We respect that the proposed framework has a better application prospect for the precise detection of URT.

Table 4 Mean location errors of different algorithms

Grid distance	Database	Test positioning target	RSSD-KNN	RSSD-WKNN	RSSD-VRKNN
1.8 m	300 MHz/20 dB	300 MHz/20 dB	1.43 m	1.35 m	1.12 m
1.8 m	300 MHz/20 dB	300 MHz/13 dB	1.52 m	1.38 m	1.17 m
1.8 m	1 GHz/20 dB	300 MHz/20 dB	1.48 m	1.32 m	1.21 m
1.8 m	1 GHz/20 dB	300 MHz/13 dB	1.45 m	1.33 m	1.15 m

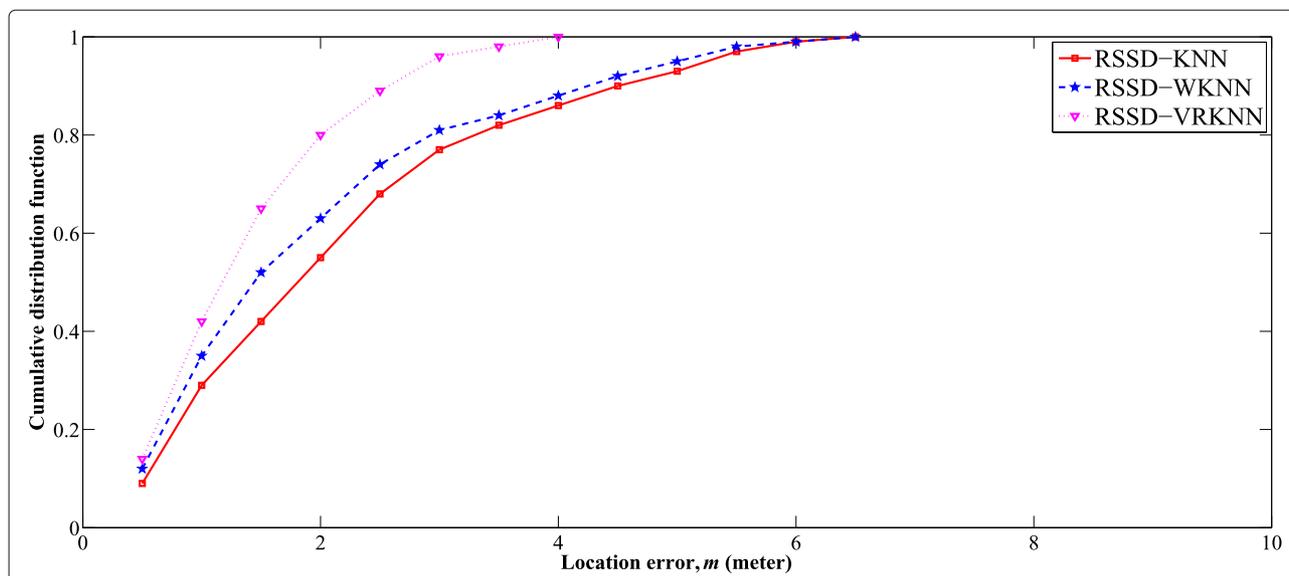


Fig. 16 CDF of location errors by different algorithms. 1 GHz/20 dB database and 300 MHz/13 dB test positioning target are utilized in our experiment

Abbreviations

AP: Access point; KNN: K -nearest neighbor; PL: Position location; RSS: Received signal strength; RSSD: Received signal strength difference; URT: Unknown radio transmitter; VR: Virtual reference

Funding

The funding was supported by Ministry of Industry and Information Technology of the People's Republic of China (CN) (No. 12-MC-KY-14).

Availability of data and materials

In our experiment, four SA44B (Signal Hound Co. Ltd.) measuring receivers as APs are utilized to collect the RSS information from the radio transmitter (TFG6300, SUIING Co. Ltd.). The experimental environment is located on the first floor of National Radio Monitoring Center, Beijing.

Authors' contributions

LZ proposed the main idea, derived the algorithm, and wrote the paper. TD review the work and versions. CJ wrote the simulation code, processed the experimental data, and revised the paper. All authors read and approved the final manuscript.

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.

Received: 19 July 2018 Accepted: 27 February 2019

Published online: 18 March 2019

References

- M. Y. Umair, K. V. Ramana, D. Yang, in *2014 IEEE Computers, Communications and IT Applications Conference*. An enhanced K -Nearest Neighbor algorithm for indoor positioning systems in a WLAN (IEEE, Beijing, 2015), pp. 19–23. <https://doi.org/10.1109/ComComAp.2014.7017163>
- A. Haniz, G. K. Tran, K. Saito, K. Sakaguchi, J. Takada, D. Hayashi, T. Yamaguchi, S. Arata, A novel phase-difference fingerprinting technique for localization of unknown emitters. *IEEE Trans. Veh. Technol.* **66**(9), 8445–8457 (2017)
- S. Chaudhari, D. Cabric, Cyclic weighted centroid algorithm for transmitter localization in the presence of interference. *IEEE Trans. Cogn. Commun. Netw.* **2**(2), 162–177 (2016)
- M. Ke, Y. Xu, A. Anpalagan, D. Liu, Y. Zhang, Distributed TOA-based positioning in wireless sensor networks: a potential game approach. *IEEE Commun. Lett.* **22**(2), 316–319 (2018)
- Z. Han, C. S. Leung, H. C. So, A. G. Constantinides, Augmented lagrange programming neural network for localization using time-difference-of-arrival measurements. *IEEE Trans. Neural Netw. Learn. Syst.* **29**(8), 3879–3884 (2018)
- A. Noroozi, A. H. Oveis, M. R. Hosseini, M. A. Sebt, Improved algebraic solution for source localization from TDOA and FDOA measurements. *IEEE Commun. Lett.* **7**(3), 352–355 (2018)
- A. A. Ababneh, Density-based sensor selection for rssi target localization. *IEEE Sensors J.* **18**(20), 8532–8540 (2018)
- D. Zhang, Y. He, X. Gong, Y. Hu, Y. Chen, B. Zeng, Multi-target AOA estimation using wideband LFM CW signal and two receiver antennas. *IEEE Trans. Veh. Technol.* **67**(8), 7101–7112 (2018)
- H. Steendam, A 3-D positioning algorithm for AOA-based VLP with an aperture-based receiver. *IEEE J. Sel. Areas Commun.* **36**(1), 23–33 (2018)
- S. Tomic, M. Beko, D. Rui, Distributed RSS-AOA based localization with unknown transmit powers. *IEEE Commun. Lett.* **5**(4), 392–395 (2016)
- Y. Wang, K. C. Ho, Unified near-field and far-field localization for AOA and hybrid AOA-TDOA positionings. *IEEE Trans. Wireless Commun.* **17**(2), 1242–1254 (2018)
- S. Tomic, M. Beko, D. Rui, 3-D target localization in wireless sensor network using RSS and AOA measurements. *IEEE Trans. Veh. Technol.* **66**(4), 3197–3210 (2017)
- J. Yin, Q. Wan, S. Yang, K. C. Ho, A simple and accurate TDOA-AOA localization method using two stations. *IEEE Signal Process. Lett.* **23**(1), 144–148 (2016)
- N. A. Jagadeesan, B. Krishnamachari, A unifying Bayesian optimization framework for radio frequency localization. *IEEE Trans. Cogn. Commun. Netw.* **4**(1), 135–145 (2018)
- M. Pauley, J. H. Manton, The existence question for maximum-likelihood estimators in time-of-arrival-based localization. *IEEE Signal Process. Lett.* **25**(9), 1354–1358 (2018)
- A. Tahat, G. Kaddoum, S. Yousefi, S. Valaee, F. Gagnon, A look at the recent wireless positioning techniques with a focus on algorithms for moving receivers. *IEEE Access.* **4**, 6652–6680 (2017)
- C. Kan, G. Ding, Q. Wu, R. Li, F. Song, Robust relative fingerprinting-based passive source localization via data cleansing. *IEEE Access.* **6**, 19255–19269 (2018)
- J. Zuo, S. Liu, H. Xia, Y. Qiao, Multi-phase fingerprint map based on interpolation for indoor localization using iBeacons. *IEEE Sensors J.* **18**(8), 3351–3359 (2018)

19. Y. Xie, Y. Wang, A. Nallanathan, L. Wang, An improved k-nearest-neighbor indoor localization method based on spearman distance. *IEEE Signal Process. Lett.* **23**(3), 351–355 (2016)
20. M. T. Van, N. V. Tuan, T. T. Son, H. Le-Minh, A. Burton, Weighted k-nearest neighbour model for indoor vlc positioning. *Let Commun.* **11**(6), 864–871 (2017)
21. M. T. Hoang, Y. Zhu, B. Yuen, T. Reese, X. Dong, T. Lu, R. Westendorp, M. Xie, A soft range limited k-nearest neighbors algorithm for indoor localization enhancement. *IEEE Sens. J.* **18**(24), 10208–10216 (2018)
22. D. Li, B. Zhang, Z. Yao, C. Li, A feature scaling based k-nearest neighbor algorithm for indoor positioning systems. *IEEE Internet Things J.* **3**(4), 590–597 (2016)
23. C. Li, Z. Qiu, C. Liu, An improved weighted k-nearest neighbor algorithm for indoor positioning. *Wirel. Pers. Commun.* **96**(2), 1–13 (2017)
24. X. Fang, Z. Jiang, L. Nan, L. Chen, Optimal weighted k-nearest neighbour algorithm for wireless sensor network fingerprint localisation in noisy environment. *Let Commun.* **12**(10), 1171–1177 (2018)
25. W. Xue, X. Hua, Q. Li, K. Yu, W. Qiu, Improved neighboring reference points selection method for wifi-based indoor localization. *IEEE Sens. Lett.* **2**(2), 2475–2472 (2018)
26. A. K. M. M. Hossain, Y. Jin, W. S. Soh, H. N. Van, Ssd: A robust RF location fingerprint addressing mobile devices' heterogeneity. *IEEE Trans. Mobile Comput.* **12**(1), 65–77 (2013)
27. A. S. El-Wakeel, A. Noureldin, H. S. Hassanein, N. Zorba, in *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)*. Utilization of Wavelet Packet Sensor De-noising for Accurate Positioning in Intelligent Road Services (IEEE, Limassol, 2018), pp. 1231–1236. <https://doi.org/10.1109/IWCMC.2018.8450418>
28. Rappaport, Theodore, *Wireless communications: principles and practice*. **8**(1), 33–8 (2001)
29. R. A. Valenzuela, in *Proceedings of Vehicular Technology Conference - VTC*. Antennas and propagation for wireless communications (IEEE, Atlanta, 2002), pp. 1–5. <https://doi.org/10.1109/VETEC.1996.503396>

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)
