IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 15, 2020, accepted February 8, 2020, date of publication February 11, 2020, date of current version February 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2973212

A Novel Weighted KNN Algorithm Based on RSS Similarity and Position Distance for Wi-Fi Fingerprint Positioning

BOYUAN WANG [1], XINGLI GAN [2,3], XUELIN LIU [1], BAOGUO YU [2,3], RUICAI JIA [2,3], LU HUANG [2,3], AND HAONAN JIA [2,3]

[1] College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China
[2] State Key Laboratory of Satellite Navigation System and Equipment Technology, Shijiazhuang 050081, China
[3] The 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang 050081, China

Corresponding author: Xingli Gan (ganxingli@163.com)

This work was supported in part by the National Key Research and Development Plan of China under Grant 2016YFB0502100 and Grant 2016YFB0502103.

ABSTRACT In Wi-Fi fingerprint positioning, what we should most care about is the distance relationship between the user and the reference points (RP). However, most of the existing weighted k-nearest neighbor (WKNN) algorithms use the Euclidean distance of received signal strengths (RSS) as distance measure for fingerprint matching, and the RSS Euclidean distance is not consistent with the position distance. To address this issue, this paper analyzes the relationship between RSS similarity and position distance, propose a novel WKNN based on signal similarity and spatial position. Firstly, we obtain the weighted Euclidean distance (WED) by balancing the size between the RSS difference and the signal propagation distance difference according to the attenuation law of the spatial signal. Then, we obtain the approximate position distance (APD) by making full use of the position distances and WEDs between RPs. Finally, the nearest RPs can be selected more accurately based on the APDs between the user and different RPs, and the position of user can be estimated by the proposed WKNN based on the APD (APD-WKNN) algorithm. In order to fully evaluate the proposed algorithm, we use three fingerprint databases for comparison experiments with eight fingerprint positioning algorithms. The results show that the proposed algorithm can significantly improve the positioning accuracy of WKNN algorithm.

INDEX TERMS Fingerprint positioning, weighted k-nearest neighbor, RSS similarity, position distance.

I. INTRODUCTION

With the rapid development of Location-based Services (LBS), many positioning technologies and signal processing methods [1]–[5] have been proposed. Due to the obstruction of the building, the usability of indoor navigation satellite signals is poor. This makes the Global Navigation Satellite System (GNSS) unable to guarantee satisfactory positioning performance in the indoor environment [6]. Therefore, various indoor positioning technologies have been proposed, among which the Wi-Fi fingerprint positioning is widely used because it can achieve positioning using only existing network facilities.

The basic idea of Wi-Fi fingerprint positioning is to use the received signal strength (RSS) of Wi-Fi signal to characterize the spatial position. The positioning process can be divided into the offline and online stages. In the offline stage, the surveyors select some points with known coordinates as reference points (RP), collect RSS from different access points (APs) at each RP, and use all RPs' RSS and coordinates to establish a fingerprint database. In the online stage, the RSS received by the user will be matched in the fingerprint database, and the user's position can be estimated based on the best matching RPs. In positioning experiments, some points with known coordinates are usually selected as test points (TP), so the positioning accuracy can be evaluated according to the estimated and actual positions of these TPs.

The method based on the nearest neighbor mechanism is most widely used in fingerprint positioning, the Weighted K-Nearest Neighbor (WKNN) [7] is the representative one. The existing WKNN algorithms generally use the Euclidean distance between RSS as the distance measure of fingerprint

The associate editor coordinating the review of this manuscript and approving it for publication was Liangtian Wan.

matching, which is called as Euclidean-WKNN in this paper. However, the RSS similarity is not equal to the closeness of the position, it is inaccurate to use RSS Euclidean distance to measure the position distance between points in space, which is mainly reflected in the following aspects:

(1) In Wi-Fi fingerprint positioning, what we should most care about is the position distance between the TP and different RPs. However, for an ideal signal environment, the differences in RSS at different positions actually reflect the difference in signal propagation distance.

(2) The free-space signal attenuation model [8] shows that the RSS attenuation and the variation of propagation distance are not a simple linear relationship.

(3) The propagation path of indoor signals is very complicated, using only RSS information to estimate the position distance will cause large positioning error. Therefore, in order to accurately describe the position distance between points, we must utilize the known position information in the positioning environment.

Therefore, this paper designs a new distance measure based on the RSS similarity and spatial position distance, and the WKNN based on approximate position distance (APD-WKNN) algorithm is proposed.

II. RELATED WORKS

In actual positioning, the RSS value will present strong fluctuations, the general approach is to collect raw RSS data during a certain time at each position, then use the average value as the offline or online RSS for fingerprint positioning. For the Euclidean-WKNN algorithm, the RSS Euclidean distance between the RP and the TP is calculated by:

$$ED_{i,*} = \sqrt{\sum_{u=1}^{M} \left(RSS_i^u - RSS_*^u\right)^2} \tag{1}$$

where $ED_{i,*}$ represents the RSS Euclidean distance between the i-th RP and the TP. The RSS_i^u and RSS_*^u are the average value of the RSS collected at the i-th RP and the TP, where the upper corner u indicates that the RSS came from the u-th AP. M is the number of APs.

Then select the RPs with the minimum RSS Euclidean distances from the TP as the nearest RPs and calculate the coordinate weights of nearest RPs. Finally, the position of the TP can be estimated by weighting the coordinates of the nearest RPs, as shown in (2) and (3).

$$w_i = \frac{1/ED_{i,*}}{\sum_{i=1}^{K} \left(1/ED_{i,*}\right)} \tag{2}$$

$$(x, y) = \sum_{i=1}^{K} w_i \cdot (x_i, y_i) \tag{3}$$

where K is the number of the nearest RPs, w_i and (x_i, y_i) are the weight and the coordinate of the i-th RP, respectively. Obviously, Euclidean-WKNN considers that a smaller RSS Euclidean distance means that the RP is closer to the TP, so it is assigned a larger weight to improve its contribution to the TP position estimation.

Niu *et al.* [9] points out that different distance measures have a significant impact on the nearest neighbor-based algorithm. Kaemarungsi and Krishnamurthy [10] also points out that using a simple RSS Euclidean distance for position estimation may cause poor positioning performance. To address this issue, many improved positioning algorithms are proposed. In [9], the experimental results show that the WKNN with Manhattan distance has better positioning performance. In [11], to deal with the noise in the Euclidean distance calculation, different weights are assigned to RSS according to their importance. Ma *et al.* [12] improves Euclidean distance by introducing the standard deviation of RSS to smooth the signal fluctuation. Niu *et al.* [13] proposes a weighted KNN method to assign different weights by defining the correlation coefficient between APs and achieve room-level positioning accuracy. Xue *et al.* [14] proposes the concept of the uneven spatial resolution of RSS and designs a weighted algorithm based on the signal attenuation model. However, the indoor Wi-Fi signal propagation is very complicated, this algorithm relies too much on the signal attenuation model and an accurate path loss exponent is difficult to obtain [15], which makes it have poor adaptability to various signal environments in practice. Bi *et al.* [16] proposes a cluster-filtered WKNN algorithm. It uses affinity propagation clustering algorithm to cluster the nearest RPs according to their position distances from each other, and the outliers are filtered out to reserve the subset with a larger number of RPs. However, the selection of RPs by this method is not based on the position distance between the RP and the TP. Therefore, some RPs close to the TP may be discarded as outliers, resulting in large positioning errors.

However, the methods mentioned above have not addressed the problem of the inconsistency between the RSS similarity and the position distance.

III. THE PROPOSED ALGORITHM

A. ANALYSIS OF THE RELATIONSHIP BETWEEN DIFFERENT DISTANCES

Taking the Euclidean distance-based WKNN as an example, the core idea of algorithm is to find the RPs with the minimum Euclidean distances of RSS, and convert the coordinates of RPs into the estimated position of the TP. We can see that the purpose of calculating the RSS Euclidean distances is to measure the position relationship between points. However, to evaluate these position distances, we can only rely on the RSS information. To make it easier to understand, we show a situation where there is a TP, two RPs and an AP. As shown in Fig.1, TP, RP_i and RP_j (the black points) represent the test point, the *i-th* RP and the *j-th* RP, respectively. PD_*, PD_i and PD_j (the black dotted lines) represent their corresponding signal propagation distances. θ_i and θ_j represent the angle between different signal propagation paths, which range from 0° to 180°. RD_i and RD_j (the black solid lines) represent the

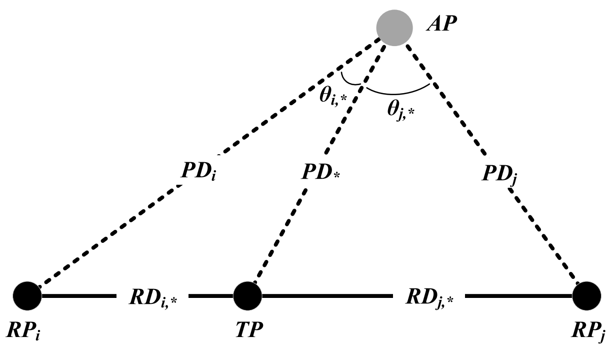


FIGURE 1. The schematic diagram of the relationship between the real position distance and the signal propagation distance.

real position distance from the TP to RPs. It should be noted that the propagation paths of indoor signal are very complex, the signal propagation distances mentioned in this paper only refer to the distances represented by linear propagation path.

We can approximately regard the real position distance and signal propagation distance as a triangular relationship. We use the triangle cosine theorem to express the position distance in terms of the propagation distance:

$$RD_{i,*} = \sqrt{PD_i^2 + PD_*^2 - 2PD_i \cdot PD_* \cdot \cos\theta_{i,*}} = \sqrt{(PD_i - PD_*)^2 + 2PD_i \cdot PD_* \cdot (1 - \cos\theta_{i,*})} \tag{4}$$

$$RD_{j,*} = \sqrt{PD_j^2 + PD_*^2 - 2PD_j \cdot PD_* \cdot \cos\theta_{j,*}} = \sqrt{\left(PD_j - PD_*\right)^2 + 2PD_j \cdot PD_* \cdot (1 - \cos\theta_{j,*})} \tag{5}$$

$$\frac{RD_{i,*}}{RD_{j,*}} = \frac{\sqrt{(PD_i - PD_*)^2 + 2PD_i \cdot PD_* \cdot (1 - \cos\theta_{i,*})}}{\sqrt{\left(PD_j - PD_*\right)^2 + 2PD_j \cdot PD_* \cdot (1 - \cos\theta_{j,*})}} \tag{6}$$

The $\theta_{i,*}$ ranges from 0° to 180°, and $(1 - \cos\theta_i)$ is greater than 0. Both $(PD_i - PD_*)^2$ and $2PD_iPD_*\,(1 - \cos\theta_i)$ are the second-order function of the propagation distance. Therefore, we assume that their contributions to RD_i are basically consistent. Moreover, in nearest neighbor-based algorithms, the proportion of the distances from different RPs to a TP, that is $RD_{i,*}/RD_{j,*}$, should be what we care about, not the specific values of distances (the values of $RD_{i,*}$ and $RD_{j,*}$). In other words, based on the proportion of different position distances, we can determine the contribution of different RPs to the position estimation. Motivated by these considerations, under the condition that the proportion of position distance does not deviate greatly, we make approximations to (6). The proportionate relationship of position distances can be described by:

$$\frac{RD_{i,*}}{RD_{j,*}} \approx \frac{\sqrt{(PD_i - PD_*)^2}}{\sqrt{\left(PD_j - PD_*\right)^2}} \cdot \frac{\sqrt{2PD_i \cdot PD_* \cdot (1 - \cos\theta_{i,*})}}{\sqrt{2PD_j \cdot PD_* \cdot (1 - \cos\theta_{j,*})}} = \frac{\triangle PD_i}{\triangle PD_j} \cdot \frac{\sqrt{2PD_i \cdot (1 - \cos\theta_{i,*})}}{\sqrt{2PD_j \cdot (1 - \cos\theta_{j,*})}} \tag{7}$$

$$\triangle PD_i = \sqrt{(PD_i - PD_*)^2} \tag{8}$$

$$\triangle PD_j = \sqrt{\left(PD_j - PD_*\right)^2} \tag{9}$$

where $\triangle PD$ represents the propagation distance difference.

The above analyses indicate that the proportion of real position distance is related to the proportion of signal propagation distance difference, and they have a non-linear relationship. Therefore, this paper considers two relationships: the relationship between the RSS and signal propagation distance difference; the relationship between the real position distance and signal propagation distance difference.

B. WEIGHTED EUCLIDEAN DISTANCE OF RSS

Now, let us analyze the relationship between the RSS and signal propagation distance difference theoretically. In Wi-Fi positioning systems, the information we can easily get from the receiver is the RSS, name and position address of the AP. Therefore, the RSS is necessary information for estimating the signal propagation distance. Many signal attenuation models are summarized in previous work [17]. For convenience, the log-distance model is adopted in this paper. As indicated in [8], it can be described by:

$$P_L\left(PD_i^u\right) = P_L\left(PD_0\right) - 10\eta log_{10}\left(\frac{PD_i^u}{PD_0}\right) \tag{10}$$

where PD_i^u represents the signal propagation distance from the u-th AP to the i-th RP. PD_0 is the reference signal propagation distance and usually set to 1 m. $P_L\left(PD_i^u\right)$ represents the RSS value of the u-th AP at the i-th RP. η is the path loss exponent, which varies with different signal propagation paths and generally ranges from 1 to 6 in indoor environment [18]. According to (10), the RSS difference can be described by:

$$\begin{aligned}\triangle RSS_{i,*}^u &= RSS_i^u - RSS_*^u \\ &= \left(P_L\left(PD_0\right) - 10\eta log_{10}\left(\frac{PD_i^u}{PD_0}\right)\right) \\ &\quad - \left(P_L\left(PD_0\right) - 10\eta log_{10}\left(\frac{PD_*^u}{PD_0}\right)\right) \\ &= -10\eta log_{10}\left(\frac{PD_i^u}{PD_0}\right) + 10\eta log_{10}\left(\frac{PD_*^u}{PD_0}\right) \\ &= -10\eta\left(log_{10}\left(PD_i^u\right) - log_{10}\left(PD_*^u\right)\right)\end{aligned} \tag{11}$$

where $\triangle RSS_{i,*}^u$ is the difference between RSS_i^u and RSS_*^u. We can see that the RSS difference has a logarithm relationship with the propagation distance.

Table 1 lists the simulated values of RSS and propagation distance calculated by the signal attenuation model. The path loss exponent η is 3, the reference distance PD_0 is 1 m and $P_L\left(PD_0\right)$ is –35 dBm. We can see that for the same size of RSS difference, the sizes of propagation distance difference under different RSS values are different. In other words, given the same size of ΔRSS, a pair of small RSS values produces a large ΔPD. This phenomenon is called as the uneven spatial resolution of RSS in [14]. As indicated in (1), the size of Euclidean distance depends only on the difference

TABLE 1. RSS value and signal propagation distance based on simulation data.

RSS (dBm)	PD (m)	ΔRSS (dBm)	ΔPD (m)
−40	1.4678		
−41	1.5849	1	0.1171
−42	1.7113	1	0.1264
−43	1.8478	1	0.1365
…	…	…	…
−61	7.3564	1	0.5435
−62	7.9433	1	0.5869
−63	8.5770	1	0.6337
…	…	…	…
−81	34.1455	1	2.5227
−82	36.8695	1	2.7240
−83	39.8107	1	2.9413

value of RSS, without considering the overall value of RSS. Therefore, the Euclidean distance cannot accurately measure the signal propagation distance difference.

To describe the difference between difference propagation distance using the RSS more accurately, we present the weighted Euclidean distance (WED). The specific approach is as follows:

For each pair of TP and RP, we use the average value to represent the overall RSS value associated with each AP, denoted by:

$$MRSS_{i,*}^{u} = \frac{1}{2}\left(RSS_i^u + RSS_*^u\right) \tag{12}$$

where $MRSS_{i,*}^{u}$ is the average RSS of the *i-th* RP and the TP, and u denotes the *u-th* AP. Then, we balance the size of RSS difference and that of Euclidean distance by assigning weights to the RSS from different APs. This weight is called as the signal weight (SW) and denoted by λ. For the *i-th* RP, the SW associated with the *u-th* AP is calculated by:

$$\lambda_{i,*}^{u} = \frac{\left(MRSS_{i,*}^{u}\right)^2}{\sum_{u=1}^{M}\left(MRSS_{i,*}^{u}\right)^2} \tag{13}$$

In this paper, the RSS readings from the receiving device are in dBm and negative. Therefore, a pair of small MRSS can produce a large SW. According to the previous analysis, we know that for the same RSS difference value, a pair of small (large) RSS values will produce a large (small) propagation distance difference. Therefore, considering the contribution of RSS differences from each AP to the size of propagation distance difference, the WED is calculated by:

$$WED_{i,*} = \sqrt{\frac{1}{M}\sum_{u=1}^{M}\lambda_{i,*}^{u}\left(RSS_i^u - RSS_*^u\right)^2} \tag{14}$$

where $WED_{i,*}$ represents the WED between the *i-th* RP and the TP. M denotes the number of the same APs detected at the RP and TP. Since the value of M is not constant at different points, to ensure the fairness of distance comparison, the WED is averaged by M. It can be concluded that the introduction of the SW can well balance the relationship between RSS and the signal distance. In other words, compared with Euclidean distance, the WED can be more accurate in describing the signal propagation distance difference.

C. THE WKNN ALGORITHM BASED ON APPROXIMATE POSITION DISTANCE

Let us continue to analyze the non-linear relationship between the real position distance and signal propagation distance difference. To our knowledge, the angle θ in (6) is difficult to obtain under the current Wi-Fi positioning systems. In addition, due to the complexity of indoor signal propagation, we cannot get a unified expression of this non-linear relationship. Therefore, the non-linear relationship should be described specifically for each pair of points. The distance weight (DW), denoted by γ, is introduced and the complex non-linear relationship of (6) is simplified into a proportional relation:

$$\frac{RD_{i,*}}{RD_{j,*}} = \frac{\triangle PD_i}{\triangle PD_j}\cdot\gamma_{i,j} \tag{15}$$

where $\gamma_{i,j}$ denotes the DW associated with the *i-th* RP, the *j-th* RP and the TP, which represents the non-linear relationship between the real position distance and the signal propagation distance difference.

Firstly, we calculate the WEDs between the TP and all RPs, and C RPs with the shortest WED are selected, called initial RPs. Then, we continue to calculate the WEDs and real position distances between initial RPs.

The WED between the *i-th* and *j-th* RPs is calculated by:

$$WED_{i,j} = \sqrt{\frac{1}{M}\sum_{u=1}^{M}\lambda_{i,j}^{u}\left(RSS_i^u - RSS_j^u\right)^2} \tag{16}$$

The real position distance is calculated by:

$$RD_{i,j} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \tag{17}$$

Now, the DW of the *i-th* RP, can be obtained by:

$$\gamma_i = \frac{1}{C-1}\sum\nolimits_{j\in\{1,\ldots,C\},j\neq i}\frac{RD_{i,j}}{WED_{i,j}} \tag{18}$$

where γ_i describes the proportional relation of the real position distance to the WED between the *i-th* RP and the other initial RPs. It can be concluded that for each initial RP, the position distance between it and the TP can be approximately estimated by its DW and the corresponding WED. Therefore, we design a new distance measure, called approximate position distance (APD), to describe the real position distance between points and enhance the WKNN algorithm. The APD between the *i-th* RP and the TP can be calculated by:

$$\begin{aligned} APD_{i,*} &= \gamma_i \cdot WED_{i,*} \\ &= \gamma_i\sqrt{\frac{1}{M}\sum_{u=1}^{M}\lambda_{i,*}^{u}\left(RSS_i^u - RSS_*^u\right)^2} \end{aligned} \tag{19}$$

We further select K RPs with the shortest APDs in the initial RP set, which are called as the nearest RPs. The coordinate weight (CW) of the nearest RPs, denoted by ω, can be calculated by:

$$\omega_i = \frac{\frac{1}{APD_{i,*}}}{\sum_{i=1}^{K} \frac{1}{APD_{i,*}}} \tag{20}$$

where ω_i is the CW of the *i-th* RP. Finally, the position of the TP can be estimated by:

$$(x, y) = \sum_{i=1}^{K} \omega_i \cdot (x_i, y_i) \tag{21}$$

The flow chart of the proposed algorithm is shown in Fig.2, and the symbol expression and definition of the main variables in this paper are also listed in Table 2.

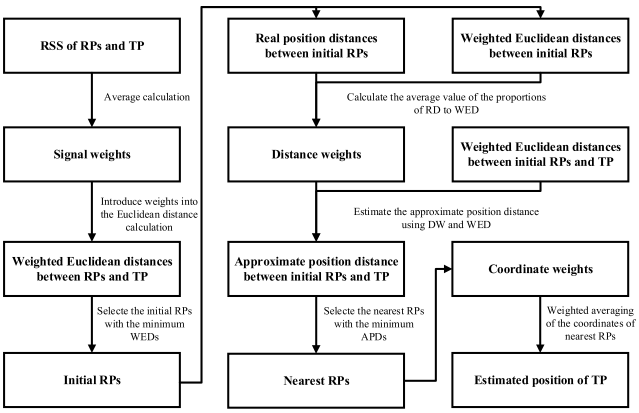


FIGURE 2. Flow chart of the proposed algorithm.

TABLE 2. The symbol and definition of main variables.

Symbol	Definition
$RD_{i,*}$	Real position distance (RD) between the *i-th* RP and the TP.
PD_i	Signal propagation distance (PD) between the *i-th* RP and an AP.
$\triangle PD_{i,j}$	Difference between PD_i and PD_j .
$ED_{i,*}$	Euclidean distance (ED) between the *i-th* RP and the TP.
$WED_{i,*}$	Weighted Euclidean distance (WED) between the *i-th* RP and the TP.
$APD_{i,*}$	Approximate position distance (APD) between the *i-th* RP and the TP.
$\lambda_{i,*}^{u}$	Signal weight (SW). It represents the RSS weight of the *u-th* AP in the calculation of the WED between the *i-th* RP and the TP.
γ_i	Distance weight (DW). It represents the proportional relation of the RD to the WED between the *i-th* RP and the other initial RPs.
ω_i	Coordinate weight (CW). It represents the contribution of the *i-th* RP to the position estimation of the TP.

To evaluate the performance of the designed distance measure APD in the ideal signal environment, we get the simulation data from the signal attenuation model and make the comparison with the Euclidean distance. For convenience, two APs, three RPs and six TPs are considered in a two-dimensional coordinate system. Table 3 lists the coordinates and RSS used for the comparison. Table 4 lists the ratios of the distances from the TP to three RPs, and the sum of the ratios equal to 1 for each TP. It should be noted that no matter which distance measure is used to measure the real position distance between location points, there will still be deviations, which is inevitable. We can see that compared with the Euclidean distance, the ratio error between APD and real position distance is less than Euclidean distance, that is, the distance ratio with APD is closer to the real position distance. The designed distance measure has better performance in measuring the position distance, which is favorable for position estimation.

IV. EXPERIMENTS AND ANALYSIS OF RESULTS

In this section, we adopt three fingerprint databases (namely Database1, Database2, Database3) and introduce eight algorithms to fully evaluate the performance of our proposed algorithm. The databases represent different RP distributions and data sizes in real indoor environments, the algorithms used for comparison can be divided into two types: the nearest neighbor-based algorithms and the machine learning-based algorithms.

A. EXPERIMENTAL DATA

For the Database1 and Database2, the Wi-Fi fingerprints are both collected on the second floor of the laboratory building in Harbin Engineering University, and the collection device is Xiaomi MIX2 Android smartphone with the Wi-Fi signal sampling frequency of 1 Hz. During the RSS collection of all RPs and TPs, the smartphone is pointing north. As shown in Fig. 3, for Database1, 10 APs are deployed, 197 points (denoted by the black solid dots) are selected as the RPs, and 109 points (denoted by the green squares) are selected as the TPs. The distance between adjacent RPs is 1.2 m, which can be considered as an intensive RP distribution. The collecting durations of each RP and TP durations of each RP and TP are 120s and 60s, respectively. For the positioning using the machine learning-based algorithms, the RPs and TPs are recorded as the training dataset and testing dataset, respectively.

To evaluate the proposed algorithm under different RP densities, we reduce the number of RPs in the Database 1 to get the Database 2. In Database2, the number of APs and the collecting duration of each point are the same as the Database1. As shown in Fig. 4, there are 100 RPs and 109 TPs in Database2. The number of RPs is nearly half that of Database 1, and the distance between RPs is also increased, which is considered as a sparse RP distribution in this paper. It should be noted that the RSS value can be influenced by the multipath effect caused by signal reflection, refraction and diffraction, as well as the signal occlusion by body [19]. Therefore, to reduce these influences, the RSS preprocessing method in [20] is adopted in the establishment

TABLE 3. The coordinates and RSS calculated by the signal attenuation model.

	AP1	AP2	RP1	RP2	RP3	
Position (m)	(2, 10)	(4, 10)	(1, 0)	(6, 0)	(11, 0)	
RSS (dBm)			(-65.06, -65.56)	(-65.97, -65.26)	(-68.87, -67.60)	
	TP1	TP2	TP3	TP4	TP5	TP6
Position (m)	(2, 0)	(3.5, 0)	(5, 0)	(7, 0)	(8.5, 0)	(10, 0)
RSS (dBm)	(-65.00, -65.26)	(-65.14, -65.02)	(-65.56, -65.06)	(-66.45, -65.56)	(-67.30, -66.20)	(-68.22, -67.00)

TABLE 4. The distance measurement comparison based on simulation data.

	Ratio of RD	Ratio of ED	Ratio of APD	Error of ED	Error of APD	Improvement
TP1	(0.07: 0.29: 0.64)	(0.05: 0.17: 0.78)	(0.09: 0.26: 0.65)	0.1855	0.0374	79.84%
TP2	(0.20: 0.20: 0.60)	(0.09: 0.14: 0.77)	(0.16: 0.21: 0.63)	0.2112	0.0510	75.85%
TP3	(0.36: 0.09: 0.55)	(0.13: 0.08: 0.79)	(0.23: 0.12: 0.65)	0.3326	0.1667	49.88%
TP4	(0.55: 0.09: 0.36)	(0.27: 0.11: 0.62)	(0.41: 0.14: 0.45)	0.3826	0.1738	54.57%
TP5	(0.60: 0.20: 0.20)	(0.38: 0.27: 0.35)	(0.49: 0.30: 0.21)	0.2753	0.1490	45.88%
TP6	(0.64: 0.29: 0.07)	(0.48: 0.40: 0.12)	(0.54: 0.39: 0.07)	0.2005	0.1414	29.48%

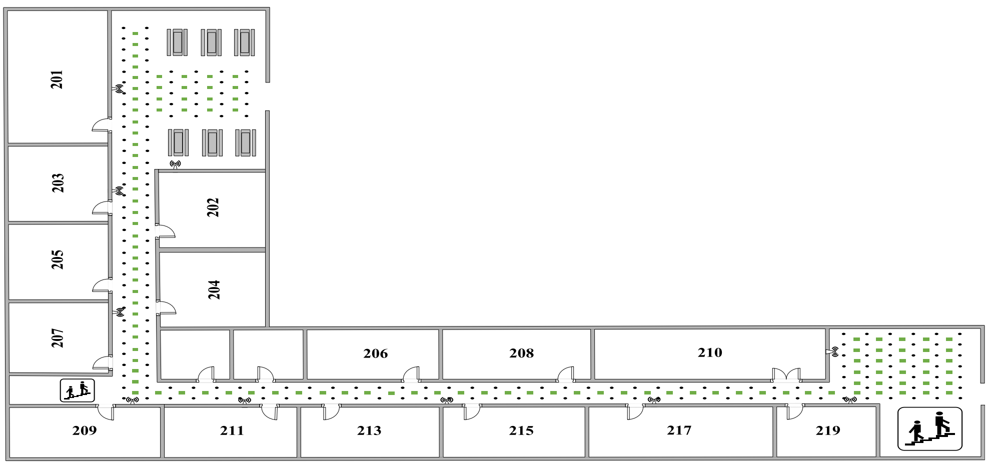


FIGURE 3. The sampling positions of Database1 with an intensive distribution of RPs.

of the Database1 and Database2. In the offline and online stages, the strongest RSS observations of each point are averaged as the RSS measurement for positioning. In this way, the processed RSS values are smoother, which can help achieve better positioning performance.

Many works have been done to solve the Wi-Fi fingerprint positioning problem. However, each method uses its own database to display the positioning results, which makes it difficult to objectively compare the performance of these methods. Therefore, to evaluate our proposed algorithm with other algorithms more fairly, the public fingerprint database UJIIndoorLoc [21] is used in the comparison experiments and called as Database3. UJIIndoorLoc was used by participants in the Evaluating Ambient Assisted Living (EvAAL) competition at IPIN 2015 [22], where the participants subjected their Wi-Fi fingerprinting solutions to a competitive benchmarking test. The database is a multi-building and multi-floor indoor fingerprint database, the data are collected by 25 different Android devices and 20 different participants in the buildings of University of Jaume I (UJI), Spain. The database consists of 19937 training samples and 1111 testing samples, which can be considered as a large database. The testing samples are taken four months later than the training ones, thus the obtained positioning error with the UJIIndoorLoc can be more realistic.

B. INFLUENCE OF THE C-VALUE ON THE PROPOSED ALGORITHM

In our proposed APD-WKNN algorithm, we select the initial RPs based on the WSD. Therefore, it is necessary to analyze the influence of the number of the initial RPs (C-value) on the performance of the proposed algorithm. To ensure the objectivity of the results, we compare the effects of different C-values given K-values on the results. It should be noted

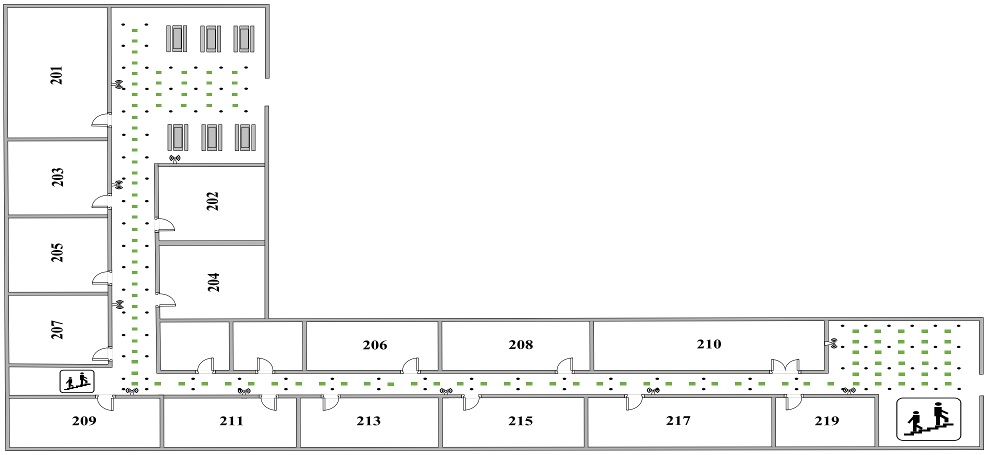


FIGURE 4. The sampling positions of Database2 with a sparse distribution of RPs.

TABLE 5. The positioning result under different K-values and C-values on Database1 and Database2.

	Database1 (K-Value = 3)		Database1 (K-Value = 4)		Database2 (K-Value = 3)		Database2 (K-Value = 4)	
C-Value	Mean (m)	RMSE (m)	Mean (m)	RMSE (m)	Mean (m)	RMSE (m)	Mean (m)	RMSE (m)
3	3.29	3.72			3.85	4.56		
4	2.75	3.03	3.16	3.49	**3.08**	3.81	3.49	3.93
5	3.05	3.50	2.51	2.97	3.14	4.15	3.77	4.45
6	3.07	3.38	2.56	2.86	3.26	4.28	**3.27**	3.68
7	2.59	2.97	**2.24**	2.60	3.30	4.33	3.38	3.98
8	**2.32**	2.78	2.39	2.81	3.19	3.94	3.32	3.90
9	2.83	3.39	2.29	3.03	3.66	4.09	3.58	4.01
10	2.44	2.81	2.57	2.83	3.82	4.23	3.79	4.72

TABLE 6. The positioning result under different K-values and C-values on Database3.

	Database3 (K-Value = 3)		Database3 (K-Value = 4)		Database3 (K-Value = 5)		Database3 (K-Value = 6)	
C-Value	Mean (m)	RMSE (m)	Mean (m)	RMSE (m)	Mean (m)	RMSE (m)	Mean (m)	RMSE (m)
3	8.27	9.85						
4	8.79	9.60	8.73	9.30				
5	8.92	9.64	6.69	7.95	7.96	9.45		
6	8.49	9.88	7.74	8.46	6.51	8.77	8.06	9.41
7	6.81	7.19	7.28	8.75	8.11	8.33	8.28	9.23
8	7.20	8.49	8.20	9.94	7.09	8.03	7.74	8.42
9	**6.59**	7.58	7.27	8.93	6.83	7.28	7.95	9.20
10	7.11	7.27	8.32	9.39	6.29	8.40	7.78	9.94
11	7.27	8.10	8.04	9.01	**6.13**	7.36	9.01	9.87
12	6.75	8.38	7.21	7.81	6.95	7.79	8.03	9.32
13	8.52	9.58	**6.74**	8.59	7.47	8.39	7.58	8.85
14	8.49	9.50	6.98	7.41	8.05	8.86	**7.11**	7.96
15	8.68	9.92	7.25	9.07	7.85	8.11	7.61	8.47
16	7.62	8.54	6.79	8.91	7.47	8.15	7.79	8.30
17	7.53	8.71	8.53	9.32	7.75	8.38	6.98	8.15
18	8.11	8.33	7.47	8.20	8.25	9.02	7.64	9.22
19	7.09	8.03	7.85	9.33	8.46	9.19	7.30	8.95
20	7.85	8.34	7.69	9.04	8.12	9.72	7.44	8.63

that for our proposed algorithm, only when C-value is not less than K-value and 2, the DWs can affect the positioning results. Considering the densities of RPs in different databases, for the Database1 and Database2, the C-values ranging from 3 to 10 are tested with the K-values from 3 to 4. For the Database3, the C-values ranging from 3 to 20 are tested with the K-values from 3 to 6. Tables 5 and 6 list the mean error and root mean square error (RMSE) under' different K-values and C-values in the Database1, Database2 and Database3. The positioning error is the straight-line distance between the real position and estimated position.

As shown in Table 5, on the condition of the K-values are 3 and 4, the average positioning error reaches the minimum of 2.32 m and 2.24 m when the C-values are 8 and 7, respectively. Because of the intensive RP distribution in Database1, a large C-value within a certain space range can provide more accurate measurement of the numerical relation between the WED and the position distance. If setting an excessively small C-value, some valid RPs near the TP may be lost. For the Database2, we can see that the average positioning error reaches the minimum of 3.08 m and 3.27 m when the C-values are 4 and 6, respectively. The distances between RPs in Database2 are larger than Database1, which can be regarded as a sparse RP distribution. If setting an oversize C-value, the coverage of initial RPs increases accordingly. This will cause that, for a TP and its nearest RPs, the calculated DWs may not reflect the relation between their WEDs and position distances well.

As shown in Table 6, for the Database3, when the K-value ranges from 3 to 6, the best positioning results are obtained with the C-values of 9, 13, 11 and 14, respectively. We know that the RSS in Database3 come from different devices and have not been preprocessed, which cause that for each pair of initial RPs, the proportions of their real position distance to the WED will vary more. Therefore, adding or subtracting an initial RP can have a greater impact on the calculation of APD and hence the positioning results.

Based on the experimental results in three databases, we can find that the effect of C-value on the positioning result is insignificant when the C-value is within a certain range. For a constant K-value, the standard deviations of the mean errors under different C-values are between 0.33 m and 0.57 m with Database1 and Database2, between 0.54 m and 0.91 m with Database3. A suitable C-value should be determined based on the site size and the RP distribution. In this paper, we provide an empirical method to set the C-value: For the intensive distribution of RPs, the C-value can be selected in the range of 1 to 3 times K-value; For the sparse distribution of RPs, the C-value can be selected in the range of 1 to 2 times K-value.

C. POSITIONING PERFORMANCE COMPARISON WITH NEAREST NEIGHBOR-BASED ALGORITHMS

In this section, we compare the positioning accuracy of the proposed APD-WKNN algorithm with four nearest neighbor-based algorithms (Euclidean-WKNN, Manhattan-WKNN, Xue *et al.* [14] and Bi *et al.* [16]) on three Databases. The Euclidean-WKNN and Manhattan-WKNN are the WKNN algorithms that use the Euclidean distance and Manhattan distance as the distance measure, respectively. Xue proposed a physical distance of RSS for indoor positioning, and the path loss exponent used in Xue's algorithm is set to 3 in the comparison experiments. Bi used the affinity propagation clustering (APC) algorithm to cluster the nearest RPs based on their position distances, and the most probable sub-cluster is reserved for position estimation by comparing the number of RPs. The positioning error in terms of the cumulative distribution function (CDF) on three databases is shown in Fig. 5, 6 and 7, respectively. Tables 7, 8 and 9 list the error statistics of these algorithms.

On Database1, the K-value is set to 3 for all algorithm, and the C-value of the proposed algorithm is set to 8. As shown in Fig. 5, the proposed algorithm obtains the best positioning

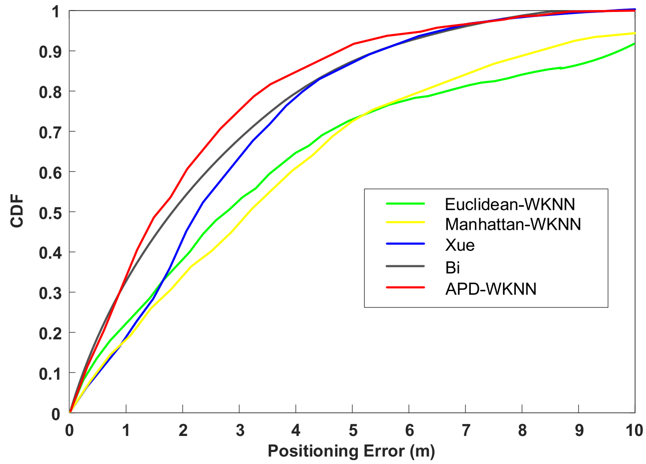


FIGURE 5. Comparison of CDF positioning errors between the proposed algorithm and related nearest neighbor-based algorithms on Database1.

accuracy. For instance, when the error threshold is 3 m and 5 m, the CDF of the proposed algorithm is 75.26% and 91.35%, which is higher than the 53.34% and 73.98% of Euclidean-WKNN, the 48.90% and 73.05% of Manhattan-WKNN, the 63.29% and 88.12% of Xue, and the 69.56% and 89.84% of Bi. As shown in Table 7, compared with other four algorithms, the mean positioning error improvements of APD-WKNN are 45.28%, 27.27%, 39.74% and 20.82%, the RMSE improvements are 48.33%, 20.80%, 36.09% and 15.24%, respectively.

TABLE 7. Positioning error statistics between the proposed algorithm and related nearest neighbor-based algorithms on Database1.

Database	Algorithm	Mean (m)	RMSE (m)
1	Euclidean-WKNN	3.85	4.35
	Manhattan-WKNN	4.24	5.38
	Xue	3.19	3.51
	Bi	2.93	3.28
	APD-WKNN	2.32	2.78

On Database2, the K-value is set to 3 for all algorithm, and the C-value of the proposed algorithm is set to 4. As shown in Fig. 6, the proposed algorithm obtains the best positioning accuracy. When the error threshold is 3 m and 5 m, the CDF of the proposed algorithm is 70.53% and 87.12%, which is higher than the 46.92% and 77.82% of Euclidean-WKNN, the 43.85% and 74.43% of Manhattan-WKNN, the 62.56% and 82.15% of Xue, and the 58.41% and 83.90% of Bi. As shown in Table 8, the mean positioning error improvements of APD-WKNN are 33.04%, 38.03%, 14.68% and 18.52%, and the RMSE improvements are 36.46%, 40.14%, 21.60% and 28.16%, respectively.

On Database3, the K-value and C-value are set to 5 and 11, respectively. As shown in Fig. 7, when the error threshold is 4 m and 6 m, the CDF of APD-WKNN is 53.94% and 70.03%, which are higher than the 42.11% and 61.25% of Euclidean-WKNN, the 34.58% and 49. 05% of Manhattan-WKNN, the 30.08% and 43.57% of Xue, and the 48.14%

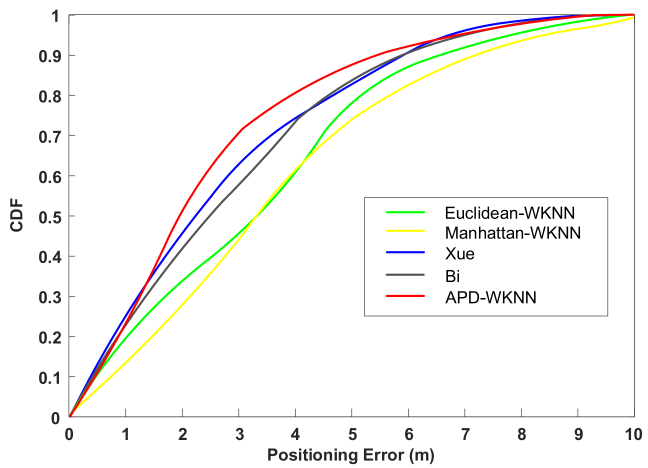


FIGURE 6. Comparison of CDF positioning errors between the proposed algorithm and related nearest neighbor-based algorithms on Database2.

TABLE 8. Positioning error statistics between the proposed algorithm and related nearest neighbor-based algorithms on Database2.

Database	Algorithm	Mean (m)	RMSE (m)
	Euclidean-WKNN	4.60	5.54
	Manhattan-WKNN	4.97	5.88
2	Xue	3.61	4.49
	Bi	3.78	4.90
	APD-WKNN	3.08	3.52

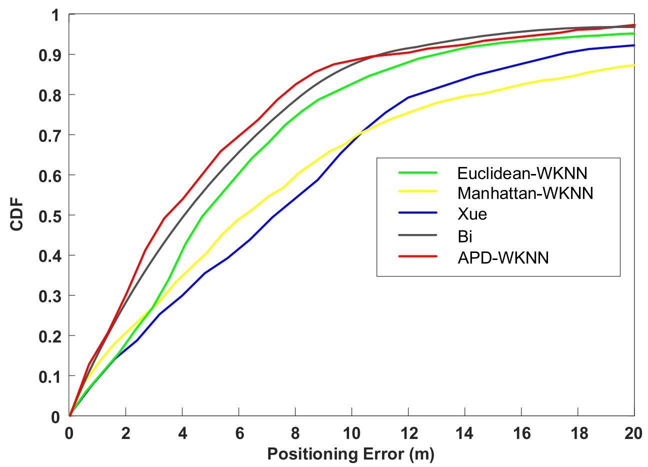


FIGURE 7. Comparison of CDF positioning errors between the proposed algorithm and related nearest neighbor-based algorithms on Database3.

and 65.31% of Bi. Xue uses the signal propagation model with fixed path loss exponent to estimate the space distance between points, while the database3 has a very large coverage, so the difference of signal propagation environment in different positions is more intense, which makes the algorithm have the worst positioning accuracy on Database3. For APD-WKNN and Bi, these two algorithms both select the RPs twice, and use the position distance between the RPs in the second selection of the nearest RPs, which can reduce the deviation of distance measurement caused by only using RSS in a large database. Therefore, we can find that the performance of APD-WKNN is slightly better than Bi, and both two algorithms outperform the others. As shown in Table 9, the mean error improvements of APD-WKNN are 20.60%, 39.67%, 44.12% and 7.40%, and the RMSE improvements are 6.61%, 35.25%, 33.86% and 3.81%, respectively.

TABLE 9. Positioning error statistics between the proposed algorithm and related nearest neighbor-based algorithms on Database3.

Database	Algorithm	Mean (m)	RMSE (m)
	Euclidean-WKNN	7.72	8.93
	Manhattan-WKNN	10.16	12.88
3	Xue	10.97	12.61
	Bi	6.62	8.67
	APD-WKNN	6.13	8.34

D. POSITIONING PERFORMANCE COMPARISON WITH MACHINE LEARNING-BASED ALGORITHMS

In addition to the classical deterministic algorithms based on the nearest neighbor mechanism, many Wi-Fi fingerprint positioning algorithms have been proposed using various machine learning methods. Therefore, to evaluate the proposed algorithm more fully, we compare the proposed algorithm with four machine learning-based algorithms (Wang *et al.* [23], Khatab *et al.* [24], Xu *et al.* [25] and Yu *et al.* [26]). Wang implemented a tree fusion-based regression model for fingerprint positioning. Khatab introduced the autoencoder (AE) to extract Wi-Fi features and used the Extreme Learning Machine (ELM) for fingerprint positioning. Xu adopted the AE for feature extraction and used the Multi-Layer Perceptron (MLP) for position estimation. Yu utilized the Radial Basis Function-based Support Vector Machine (RBF-SVM) for fingerprint positioning and achieved high accuracy. The parameters of these algorithms are set to the same as them in the literatures, and no changes are made in this paper. The K-values and C-values are also the same as the previous section. The positioning error in terms of CDF is shown in Fig. 8, 9 and 10, respectively. Tables 10, 11 and 12 list the positioning error statistics.

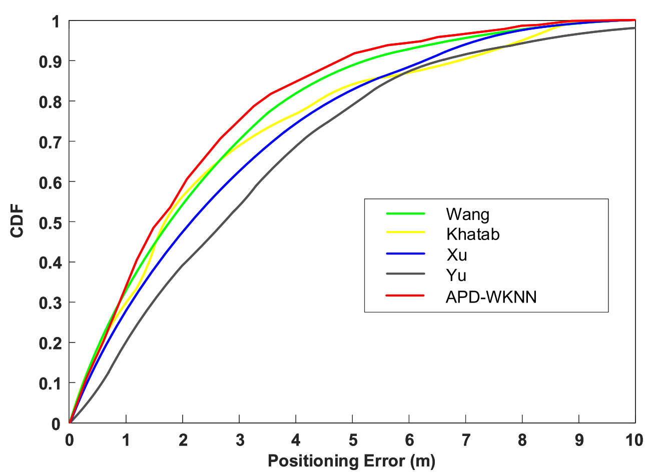


FIGURE 8. Comparison of CDF positioning errors between the proposed algorithm and machine learning-based algorithms on Database1.

The positioning error on Database1 is shown in Fig. 8 and Table 10. The 50%, 75% and mean error of the APD-WKNN are 1.58 m, 3.06 m and 2.32 m, which are smaller than 1.79 m, 3.34 m and 2.74 m of Wang, 1.70 m, 3.83 m and 2.98 m of Khatab, 2.15 m, 4.09 m and 3.35 m of Xu, 2.75 m, 4.59 m and 3.89 m of Yu. Compared with the listed neighbor-based algorithms, the mean error improvements of the APD-WKNN on Database1 are 15.33%, 22.15%, 30.75% and 40.36%, and the RMSE improvements are 9.15%, 18.71%, 26.65% and 34.74%, respectively.

TABLE 10. Positioning error statistics between the proposed algorithm and machine learning-based algorithms on Database1.

Database	Algorithm	Mean (m)	RMSE (m)
1	Wang	2.74	3.06
	Khatab	2.98	3.42
	Xu	3.35	3.79
	Yu	3.89	4.26
	APD-WKNN	2.32	2.78

The positioning error on Database2 is shown in Fig. 9 and Table 11. The 50%, 75% and mean error of the APD-WKNN are 1.97 m, 3.37 m and 3.08 m, which are smaller than 2.40 m, 3.90 m and 3.53 m of Wang, 2.29 m, 4.16 m and 3.96 m of Khatab, 2.78 m, 4.49 m and 4.33 m of Xu, 2.71 m, 5.19 m and 4.74 m of Yu. Compared with the listed neighbor-based algorithms, the mean error improvements of the APD-WKNN on Database2 are 12.75%, 22.23%, 28.87% and 35.02%, and the RMSE improvements are 9.51%, 19.08%, 28.74% and 32.44%, respectively.

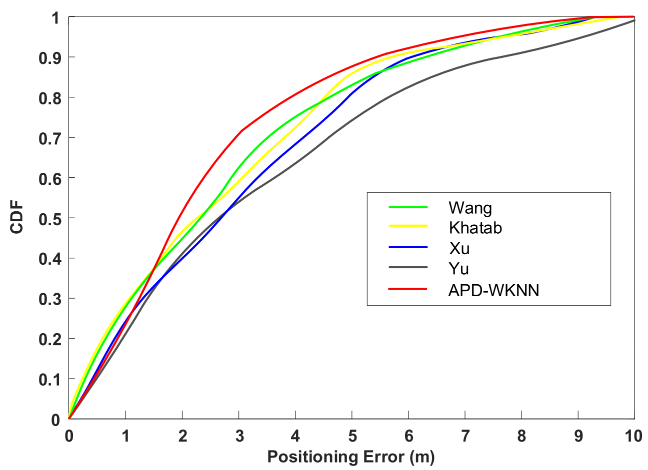


FIGURE 9. Comparison of CDF positioning errors between the proposed algorithm and machine learning-based algorithms on Database2.

TABLE 11. Positioning error statistics between the proposed algorithm and machine learning-based algorithms on Database2.

Database	Algorithm	Mean (m)	RMSE (m)
2	Wang	3.53	3.89
	Khatab	3.96	4.35
	Xu	4.33	4.94
	Yu	4.74	5.21
	APD-WKNN	3.08	3.52

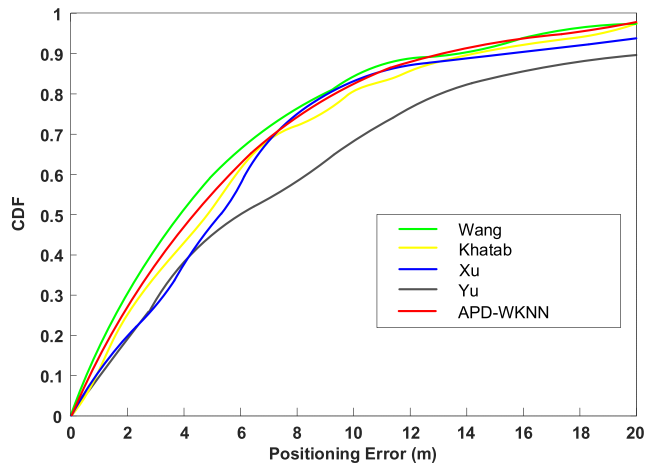


FIGURE 10. Comparison of CDF positioning errors between the proposed algorithm and machine learning-based algorithms on Database3.

TABLE 12. Positioning error statistics between the proposed algorithm and machine learning-based algorithms on Database3.

Database	Algorithm	Mean (m)	RMSE (m)
3	Wang	6.32	7.59
	Khatab	7.23	9.04
	Xu	7.85	9.73
	Yu	10.64	11.81
	APD-WKNN	7.05	8.27

The positioning error on Database3 is shown in Fig. 10 and Table 12. We can find that our proposed algorithm, Wang's and Khatab's algorithms obtain similar positioning accuracy. The 50% errors of the three algorithms are 4.30 m, 3.93 m, 4.85 m, and the 75% errors are 8.22 m, 7.81 m, 8.89 m. Their mean errors are 7.05 m, 6.32 m and 7.23 m, respectively. Among them, Wang's algorithm has the highest average positioning accuracy. The method with the lowest accuracy is still Yu's, and its average positioning error is only 10.64 m. We can see that the positioning accuracy with Database3 is lower than Database1 and Database2, this is because the sampling data of Database3 are the original data without averaging or other preprocessing, collected by 25 different devices and 20 different participants, this data feature makes it more challenging to use nearest neighbor-based algorithm for positioning. For machine learning-based algorithms, the fingerprint feature extraction can reduce the influence of heterogeneous devices' RSS inconsistency on positioning. Besides, the powerful classification and fitting capabilities of machine learning enable them to cope with such large database. Nevertheless, we can see that, by using the more effective distance measure and accurate weigh, the mean positioning error of the proposed algorithm is still at a low level, which is 10.19% less than Xu and 33.74% less than Yu.

Moreover, compared with the positioning methods based on machine learning, our proposed algorithm has no training process and does not need adjusting numerous parameters.

Therefore, the complexity of the proposed algorithm is lower than that of machine learning-based algorithms, which is very important for indoor positioning based on mobile devices such as smartphones.

V. CONCLUSION

This paper proposes a novel WKNN algorithm based on a new distance measure for Wi-Fi fingerprint positioning. We first analyze the relationship between RSS similarity and signal propagation distance difference, and present the weighted Euclidean distance based on the attenuation law of the spatial signal. Then, by combining the weighted Euclidean distance with the known position information of RPs, the approximate position distance is designed and used for improving the WKNN algorithm. The positioning experiments are conducted with three databases, and the proposed algorithm is compared with eight different types of algorithms to evaluate the performance. The results show that the designed approximate position distance outperforms other distance measures in describing the position relation of points. The mean positioning accuracy of the proposed algorithm outperforms the Euclidean-WKNN by 45.28%, 38.41% and 20.60%, outperforms the SVM-based algorithm by 40.36%, 37.94% and 33.74% in three databases, respectively. In addition to the K-value and C-value, the algorithm does not need to set additional parameters, and the effect of C-value on the positioning result is insignificant when the C-value is within a certain range. In future work, we will focus on the automatic construction and update of database to reduce the work of offline field survey.

REFERENCES

[1] Z. Li, J. Shi, X. Wang, and F. Wen, "Joint angle and frequency estimation using one-bit measurements," *Sensors*, vol. 19, no. 24, p. 5422, Dec. 2019.

[2] L. Wan, X. Kong, and F. Xia, "Joint range-Doppler-angle estimation for intelligent tracking of moving aerial targets," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1625–1636, Jun. 2018.

[3] L. Wan, L. Sun, X. Kong, Y. Yuan, K. Sun, and F. Xia, "Task-driven resource assignment in mobile edge computing exploiting evolutionary computation," *IEEE Wireless Commun.*, vol. 26, no. 6, pp. 94–101, Dec. 2019.

[4] H. Wang, L. Wan, M. Dong, K. Ota, and X. Wang, "Assistant vehicle localization based on three collaborative base stations via SBL-based robust DOA estimation," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5766–5777, Jun. 2019.

[5] X. Wang, L. Wan, M. Huang, C. Shen, and K. Zhang, "Polarization channel estimation for circular and non-circular signals in massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1001–1016, Sep. 2019.

[6] M. Susi, V. Renaudin, and G. Lachapelle, "Motion mode recognition and step detection algorithms for mobile phone users," *Sensors*, vol. 13, no. 2, pp. 1539–1562, Jan. 2013.

[7] X. Fang, Z. Jiang, L. Nan, and L. Chen, "Optimal weighted K-nearest neighbour algorithm for wireless sensor network fingerprint localisation in noisy environment," *IET Commun.*, vol. 12, no. 10, pp. 1171–1177, Jun. 2018.

[8] X. Liu, S. Zhang, Q. Zhao, and X. Lin, "A novel approach for fingerprint positioning based on spatial diversity," in *Proc. 3rd Int. Conf. Adv. Comput. Theory Eng. (ICACTE)*, vol. 6, Aug. 2010, pp. V6-441–V6-445.

[9] J. Niu, B. Lu, L. Cheng, Y. Gu, and L. Shu, "ZiLoc: Energy efficient WiFi fingerprint-based localization with low-power radio," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2013, pp. 4558–4563.

[10] K. Kaemarungsi and P. Krishnamurthy, "Properties of indoor received signal strength for WLAN location fingerprinting," in *Proc. Int. Conf. Mobile Ubiquitous Syst., Netw. Services*, Nov. 2004, pp. 14–23.

[11] Y. Gao, J. Niu, R. Zhou, and G. Xing, "ZiFind: Exploiting cross-technology interference signatures for energy-efficient indoor localization," in *Proc. Int. Conf. Comput. Commun.*, Apr. 2013, pp. 2940–2948.

[12] R. Ma, Q. Guo, C. Hu, and J. Xue, "An improved WiFi indoor positioning algorithm by weighted fusion," *Sensors*, vol. 15, no. 9, pp. 21824–21843, Sep. 2015.

[13] J. Niu, B. Wang, L. Shu, T. Q. Duong, and Y. Chen, "ZIL: An energy-efficient indoor localization system using ZigBee radio to detect WiFi fingerprints," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1431–1442, Jul. 2015.

[14] W. Xue, X. Hua, Q. Li, K. Yu, W. Qiu, B. Zhou, and K. Cheng, "A new weighted algorithm based on the uneven spatial resolution of RSSI for indoor localization," *IEEE Access*, vol. 6, pp. 26588–26595, 2018.

[15] K. Kaemarungsi and P. Krishnamurthy, "Modeling of indoor positioning systems based on location fingerprinting," in *Proc. Int. Conf. Comput. Commun.*, vol. 2, Feb. 2004, pp. 1012–1022.

[16] J. Bi, Y. Wang, X. Li, H. Qi, H. Cao, and S. Xu, "An adaptive weighted KNN positioning method based on omnidirectional fingerprint database and twice affinity propagation clustering," *Sensors*, vol. 18, no. 8, p. 2502, Aug. 2018.

[17] R. G. Akl, D. Tummala, and X. Li, "Indoor propagation modeling at 2.4 GHz for IEEE 802.11 networks," in *Proc. IASTED Int. Multi Conf. Wireless Opt. Commun.*, Jul. 2006, pp. 12–17.

[18] S. Y. Seidel and T. S. Rappaport, "914 MHz path loss prediction models for indoor wireless communications in multifloored buildings," *IEEE Trans. Antennas Propag.*, vol. 40, no. 2, pp. 207–217, Feb. 1992.

[19] I. Alshami, N. Ahmad, S. Sahibuddin, and F. Firdaus, "Adaptive indoor positioning model based on WLAN-fingerprinting for dynamic and multi-floor environments," *Sensors*, vol. 17, no. 8, p. 1789, Aug. 2017.

[20] W. Xue, W. Qiu, X. Hua, and K. Yu, "Improved Wi-Fi RSSI measurement for indoor localization," *IEEE Sensors J.*, vol. 17, no. 7, pp. 2224–2230, Apr. 2017.

[21] J. Torres-Sospedra, R. Montoliu, A. Martinez-Uso, J. P. Avariento, T. J. Arnau, M. Benedito-Bordonau, and J. Huerta, "UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Oct. 2014, pp. 261–270.

[22] A. Moreira, M. J. Nicolau, F. Meneses, and A. Costa, "Wi-Fi fingerprinting in the real world–RTLS@UM at the EvAAL competition," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Oct. 2015, pp. 1–10.

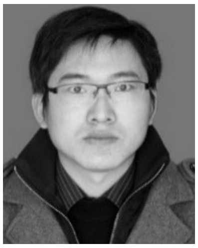
[23] R. Wang, Z. Li, H. Luo, F. Zhao, W. Shao, and Q. Wang, "A robust Wi-Fi fingerprint positioning algorithm using stacked denoising autoencoder and multi-layer perceptron," *Remote Sens.*, vol. 11, no. 11, p. 1293, May 2019.

[24] Z. E. Khatab, A. Hajihoseini, and S. A. Ghorashi, "A fingerprint method for indoor localization using autoencoder based deep extreme learning machine," *IEEE Sens. Lett.*, vol. 2, no. 1, pp. 1–4, Mar. 2018.

[25] C. Xu, Z. Jia, P. Chen, and B. Wang, "CSI-based autoencoder classification for Wi-Fi indoor localization," in *Proc. Chin. Control Decis. Conf. (CCDC)*, May 2016, pp. 6523–6528.

[26] F. Yu, M. H. Jiang, J. Liang, X. Qin, M. Hu, T. Peng, and X. R. Hu, "An indoor localization of WiFi based on support vector machines," *Adv. Mater. Res.*, vols. 926–930, pp. 2438–2441, May 2014.

BOYUAN WANG was born in Inner Mongolia, China, in 1992. He received the B.E. degree in electronics and information engineering from Harbin Engineering University, Heilongjiang, China, in 2013, where he is currently pursuing the Ph.D. degree. He participated in the National Key Research and Development Plan of China and published many articles, including two articles indexed by the Science Citation Index and four articles indexed by the EI Compendex. His research interests include indoor positioning, machine learning, pattern recognition, and signal processing.

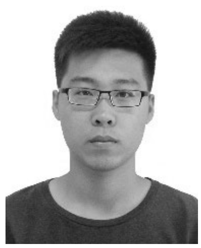


XINGLI GAN was born in Tangshan, Hebei, China, in 1981. He received the Ph.D. degree from Harbin Engineering University, China, in 2008. He is currently a Professor with the State Key Laboratory of Satellite Navigation System and Equipment Technology. He is the author of one book, more than 40 articles, and more than 15 inventions. His research interests include satellite navigation, indoor pseudo-satellite positioning, visual positioning, artificial intelligence, and location service application.

XUELIN LIU was born in 1965. He received the B.E. degree in wireless communication engineering from the Changchun Institute of Posts and Telecommunications, China, in 1986, the M.E. degree in communication and electronic systems from the Shijiazhuang Communication and Measurement Technology Research Institute, China, in 1989, and the Ph.D. degree from the College of Management and Economics, Tianjin University, China, in 2007. He is currently a Professor and a Ph.D. Supervisor with the College of Information and Communication Engineering, Harbin Engineering University. He is also the Chairman of Taiji Computer Corporation. His research interests include sensor technology, indoor positioning, cloud computing, the Internet of Things, and wireless communication.

BAOGUO YU was born in 1966. He received the B.E. degree in information processing from the National University of Defense Technology, China, in 1988, the M.E. degree in circuit and system from Yanshan University, China, in 1995, and the Ph.D. degree in communication and information system from the Beijing Institute of Technology, China, in 2005. He is a Professor and a Ph.D. Supervisor, and the Chief Scientist of China Electronic Technology Group Corporation. He is currently the Director of the State Key Laboratory of Satellite Navigation System and Equipment Technology. His research interests include GNSS, indoor positioning, applications of multisensor systems, and signal processing. He is a Fellow of Chinese Institute of Electronics (CIE). He has received many national science and technology awards.

RUICAI JIA was born in Linyi, Shandong, China, in 1986. He received the Ph.D. degree from Harbin Engineering University, China, in 2012. He is currently a Senior Engineer with China Electronic Technology Group Corporation. His research interests include integrated navigation, inertial navigation, and pedestrian navigation. He is a Fellow of CIE.



LU HUANG was born in Panjin, Liaoning, China, in 1991. He received the B.E. degree in electronic information engineering from Shenyang Ligong University, in 2014, and the M.E. degree in information and communication engineering from Harbin Engineering University, in 2017. His research interests include indoor multisensor fusion positioning and indoor pseudo-satellite positioning. He participated in five national key Research and Development projects, published more than ten academic articles, and applied for three invention patents.

HAONAN JIA was born in Handan, Hebei, China, in 1990. He received the B.E. and M.E. degrees from Harbin Engineering University, Heilongjiang, China, in 2014 and 2017, respectively. His research interests include indoor positioning, inertial navigation, and pedestrian navigation.