Received May 8, 2019, accepted May 21, 2019, date of publication May 27, 2019, date of current version June 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2919329

Learning to Improve WLAN Indoor Positioning Accuracy Based on DBSCAN-KRF Algorithm From RSS Fingerprint Data

KAI WANG[1], XING YU[1], QINGYU XIONG[2], QIWU ZHU[1], WANG LU[1], YA HUANG[1], AND LINYU ZHAO[1]

[1]School of Automation, Chongqing University, Chongqing 400044, China
[2]School of Big Date & Software Engineering, Chongqing University, Chongqing 400044, China

Corresponding author: Kai Wang (akyle@163.com)

This work was supported in part by the Chongqing Research Program of Basic Research and Frontier Technology under Grant cstc2017jcyjBX0025, in part by the Science and Technology Major Special Project of Guangxi under Grant GKAA17129002, and in part by the Fundamental Research Funds for the Central Universities under Grant 106112017CDJXY170005.

ABSTRACT WLAN-based indoor positioning algorithm has the characteristics of simple layout and low price, and it has gradually become a hotspot in both academia and industry. However, due to the poor stability of Wi-Fi signals, received signal strength (RSS) fingerprints of some adjacent reference positions are difficult to evaluate similarity when utilizing traditional distance-based calculation methods. By clustering these RSS fingerprints into one region, the commonly utilized KNN algorithm in the past cannot achieve accurate positioning in the region. For this, we introduce a concept of the insensitive region of the RSS fingerprint and a new algorithm named DBSCAN-KRF. This algorithm can delete noise sample and detect insensitive region, then, different methods are selected to achieve indoor positioning by judging the region of the estimated fingerprint sample, the KNN algorithm is selected when the region is sensitive, and random forest algorithm is selected when the region is insensitive. The experimental results show that the DBSCAN-KRF algorithm is superior while compared with other alternative indoor positioning algorithms.

INDEX TERMS WLAN indoor position, control and optimization, machine learning, DBSCAN-KRF algorithm, fingerprint data.

I. INTRODUCTION

In the wake of the rapid development of wireless communication technology and computer technology, location-based service(LBS) has received much attention from both academia and industry, the most critical issue for it is how to use positioning technology to obtain current position accurately and rapidly. According to different application scenarios, positioning technology can be divided into outdoor positioning and indoor positioning. The outdoor positioning technology mainly relies on the traditional satellite positioning, such as Global Positioning System (GPS) and BeiDou Navigation Satellite System (BDS), these systems are very mature and have been continuously optimized, they can basically meet the needs of people in the outdoor environment. However, people require higher-accuracy positioning in indoor environment, and the GPS signal may be obstructed by the building wall, so the satellite positioning technology cannot provide better indoor positioning effect [1]. At present, wireless communication technology and image processing technology are widely used to achieve indoor positioning. Among them, the accuracy of image processing technology is very high, especially when the deep learning method is applied to the image processing. However, it must use camera equipment frequently to take pictures, the inconvenience of this technology makes it difficult to apply in specific applications [2], [3]. Wireless communication technology has many advantages, such as convenient installation, high maturity, low cost and small size, it has been widely applied in indoor positioning [4].

In the indoor environment, wireless communication technology is commonly used to achieve positioning, such as ultra-wideband (UWB), Zigbee, geomagnetism or WLAN and other technologies [5]. Among them, the accuracy of

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

UWB and Zigbee technology is very high, they can achieve positioning with meter level accuracy, but the equipment is expensive, they are not suitable for large-area indoor environment [6], [7]. The geomagnetic technology is much cheaper than ultra-wideband or Zigbee, but its accuracy can only reach 100 meters or so [8]. Therefore, all of the above methods are not suitable for large-scale smart building. In recent years, while moving ahead with the smart-cities, the coverage area of WLAN network inside the building become more and more wide, therefore, using WLAN technology to achieve indoor positioning can avoid the purchase of special signal transmission and receiving equipment, it can lead to greatest reduction in cost [9]. Moreover, the accuracy of WLAN positioning can reach 1-10 meters by using appropriate positioning algorithm, which basically meets the needs of indoor positioning [10].

The method of indoor positioning based on WLAN signal strength can be divided into two categories, the triangle algorithm and position fingerprint algorithm. The triangulation algorithm can estimate the position by using the distance information between the measured point and the reference points. Its high accuracy strongly relies on the certain AP position and the accurate loss model for Wireless signal transmit, but in fact, due to the complexity of indoor environment, it is very hard to establish accurate model, so this method is rarely used in reality [11]. On the contrary, the position fingerprint algorithm acquires the position by comparing the fingerprint information of estimated point with that in fingerprint library. Position fingerprint is a set of RSS values received from various APs at a certain time and position, it is a purely data-driven method, hence we need not set the accurate loss model for Wireless signal transmit, these advantages make the method prevalent in the area of indoor positioning under complex situation [12], [13].

II. RELATE WORK

A. SYSTEM ARCHITECTURE

In general, the fingerprint-based indoor positioning system can be divided into two phases of operation, online and offline, just like the below picture 1 [14].

As shown in Figure 1. In offline phase, the sign information of each Reference Points(RPs) should be gathered to build a fingerprint database, we can obtain a position model by training these fingerprint data, this model can reflect the mapping relationship between signal information and position. This positioning model will be used in the online phase, in this phase, the current position can be estimated by entering the instantaneous signal into the positioning model.

B. BUILD FINGERPRINT DATABASE

Several RPs are set evenly in the positioning area, at each RP, mobile terminal can be used to collecting signal strength data from each access point(AP) for several times, this data is the fingerprint information of the RP [15]. After storing the fingerprint information of all the reference points in a database, the establishment of fingerprint database is completed.

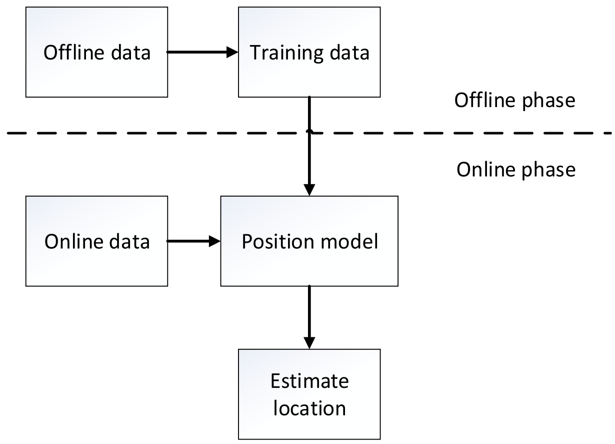


FIGURE 1. The framework of fingerprint-based indoor positioning system.

In fingerprint database, each received signal strength (RSS) fingerprint consists of two parts, one of which is physical position of RPs in the positioning area, and the other is the signal strength data received from APs.

C. POSITIONING MODEL

A high-accuracy positioning system depend on a good positioning model, and the model is based on the fingerprint data in the database, how to deal with these data is the key to reducing positioning error and improve accuracy [16].

For this, a positioning algorithm based on K-nearest neighbor, pioneered by Bahl et al. This method can find out the K nearest positions by calculating the distance between the measured-position fingerprint and all the fingerprints in the database, the target position is classified by a majority vote of its K nearest positions, with the position being assigned to the class most common among its k nearest positions. It is relatively simple and easy to implement, but owing to each positioning needs to calculate the distance too many times, the time complexity of this algorithm is very high, especially when the positioning area is large. Furthermore, it has no anti-interference ability and may cause a great error under complex situation.

In order to improve positioning speed, Zheng Wu proposed a positioning method using online independent support vector machine (OISVM) and undersampling techniques. Compared with traditional SVMs, this method removes the borderline samples and optimizes the kernel parameter. The time complexity of both the training process and the prediction process are significantly decreased [17]. Jun Ma proposed a scheme called "cluster filtered KNN" (CFK). This method utilizes clustering technique to partition all the nearest K neighbors into different clusters. In the end, the final estimated position can be calculated based on the elements of only one cluster selected from them. The algorithm improves positioning accuracy effectively through such filtering work. However, the clustering technique in this algorithm is only used for the final position calculation, and the original fingerprint signals are not clustered and analyzed [18].

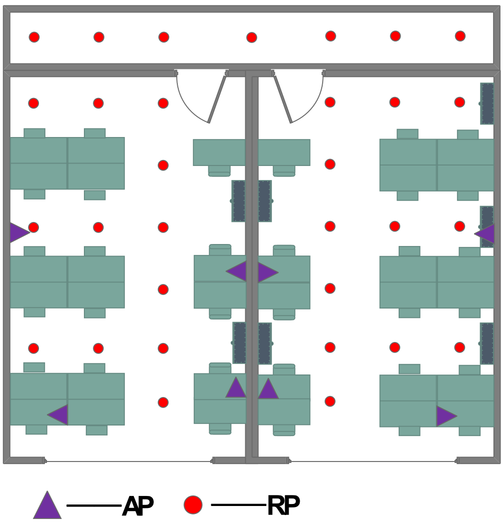


FIGURE 2. Floor plan of indoor positioning experiment.

D. DBSCAN

DBSCAN (Density-Based Spatial Clustering of Appli-cations with Noise) is a density-based spatial clustering algorithm [19]. This algorithm can group together very close points as high-density region clusters, the shape of it could be arbitrarily, hence it can adapt to a variety of data sets. Besides that, DBSCAN has a notion of noise, it can mark points in those low-density region as outliers, so it is robust to noise. It can deal with noise points effectively and discover arbitrary shapes of spatial clustering.

III. DATA ACQUISTION

In order to better reflect the performance of each algorithm more comprehensively, choosing a real environment to test, all of the data comes from the real environment.

A. DIVIDE THE POSITIONING AREA

Taking into account a number of factors, 2 offices and corridors in 25th floor of teaching building is considered as an experimental area, the 2 offices have the same pattern and are separated by a brick wall, each of them is a 8.8m $\times$ 5.6m room. These rooms have many desks and cupboards, and there is a constant flow of people in it, so it is a complex environment, the RSS data collected under this situation is representative.

The area of room is about 49 square meters, according to the principle that each RPs should be separated by 1.5m, the layout of our testbed can be divided into around 18 RPs, however, because there are some RPs placed on the desk, so actually each room has 12 reference points. The entire area contains 2 offices and corridors, there are 31 reference points in all.

The floor plan with RPs and APs as shown in Figure 2.

B. RETRIEVE DATA

The effective distribution of WLAN signal intensity indoo-rs is sensitive to several factors, besides the changeless indoor objects, the movable human body also have an impact on it. The human body is a very important interference source in WLAN signal propagation, whether the fingerprint data collection in offline phase, or the real-time signal reception in online phase, both the two process require the participa-tion of people [20]. In reality, walking is a very common pheno-menon in indoor environment, so we must consider the factor of human body in the preliminary test and practical application.

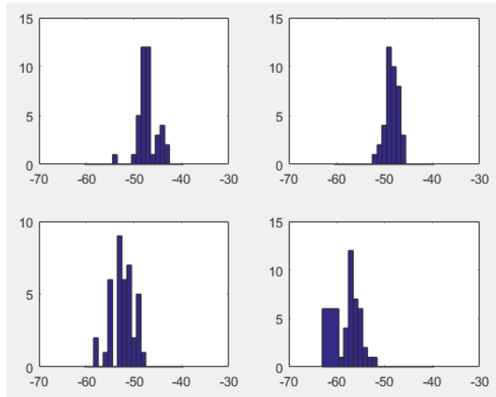


FIGURE 3. Histogram of the signal intensity in four directions.

The Figure 3 shows the signal intensity histogram from the same AP at the same RP position when the people hand-held receiver and face four directions. It can be clearly seen that the signal intensity varies greatly in different directions. Therefore, measurements in a single direction can lead to absence of an effective fingerprint. In order to obtain more comprehensive signal, a reference point should be divided into four directions, and collect 10 groups of fingerprints on each direction.

IV. PROPOSED METHODS

A. DBSCAN ALGORITHM

In positioning system, if the positioning area is too large, we need to collect a mass amount of RSS samples to achieve better positioning effect, it is inevitable that there will be a certain amount of noise samples in these raw data. Therefore, it would waste large computing resources and time resources by using these raw data directly, even will severely reduce the accuracy of positioning [21], [22]. In order to solve these issues, the clustering algorithm should be introduced to divided all RSS samples into clusters [23], [24].

According to the similarity of data samples, clustering analysis can classify the data and find the distribution characteristics of object space. There are many clustering algorithms have been proposed, such as K-means, SOM, Hierarchical Clustering and DBSCAN etc [25], [26]. Among them, DBSCAN algorithm is a density-based clustering algorithm, it can not only discover arbitrarily shaped clusters, but also have the ability to handle outliers effectively. There-fore, this paper adopts DBSCAN algorithm as clustering algorithm.

Every point in the dataset could be divided into three kinds by using DBSCAN method, core point, border point and

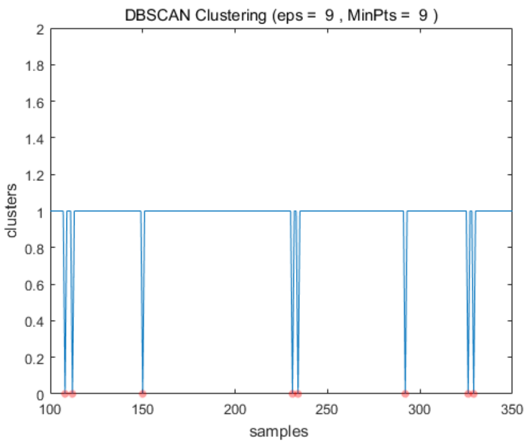


FIGURE 4. DBSCAN clustering result under the 9 eps and 9 minpts.

noise point [27]. Each point belongs to what kind of point depends on the sample distribution and two input parameters eps and minpts of DBSCAN algorithm. The former parameter is the neighbor radius when defining the density, if the radius region contains sufficiently many points, a cluster is started. The latter parameter is the minimum number of points required to form a dense region, the type of point is determined by comparing the number of points in the radius region with the parameter minpts. After determining the specific parameters, a random unvisited point x will be selected and count the number of points in its eps-neighborhood N_x, if the density $|N_x| \leq minpts$, marked this point as noise point, otherwise, it is considered as core point and a new cluster is created. Then in accordance with the above principles, calculate the density of each point and identify the core points in this eps-neighborhood of the cluster, if its density more than minpts, marked it as core point and expand the region of the cluster by assigning it to this cluster, if less than minpts, it is considered as a border point. When there are no unvisited points can be assigned to cluster, the new cluster has been completed. Repeating the above clustering process until all the points are assigned to some cluster or marked as noise point, the process is halted and the clustering process is complete.

In this study, using this method to complete the noise processing and insensitive points processing. From the above principle, we can see that the distance calculation method in DBSCAN algorithm is Euclidean distance, which is also used in KNN algorithm. Therefore, the noise points detected by DBSCAN must affect the positioning accuracy, we need to detect and delete it in the data preprocessing. When setting a larger eps and minpts parameters, the clustering process usually gets only one cluster, we can mark the outliers as noise sample in all samples.

The below figure 4 shows the DBSCAN clustering result under the 9 eps and 9 minpts, these outliers are labeled cluster 0 and other sample are labeled cluster 1, it should be clear that there are some samples are the noise sample, sample 292 is one of them.

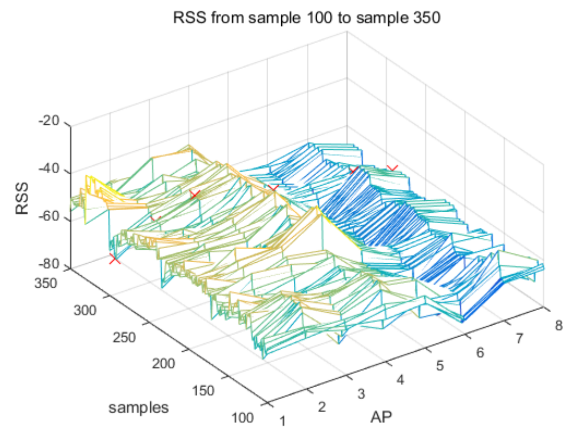


FIGURE 5. RSS values from sample 100 to sample 350.

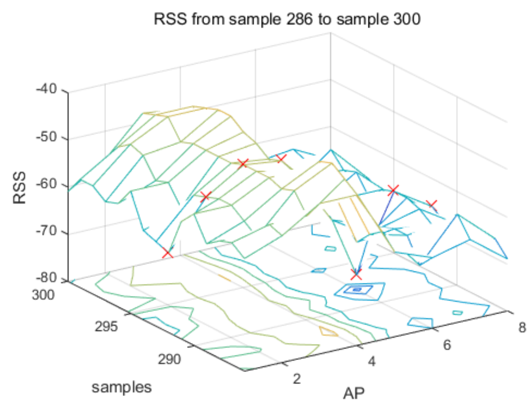


FIGURE 6. RSS values from sample 286 to sample 300.

The figure 5 shows the RSS values from sample 100 to sample 350, each sample has eight dimensions which stand for the strength of RSS from different APs. Among them, the adjacent samples come from the same RP and direction, it can be seen that the similarity of the adjacent samples is relatively high. There are still individual noise points which perform differently, the marked points in Fig.5 are different dimensions of the noise sample found by DBSCAN algorithm. Due to the large number of samples contained in this picture, it not very easy to find the abnormal point 292 which is marked in the chart.

In order to show the outliers clearly, the local of the picture is enlarged and contain 15 samples, as shown in figure 6, sample 292 is marked as an outlier.

Besides the outliers, insensitive points can also reduce the positioning accuracy. When setting a small eps and minpts parameters, DBSCAN clustering can get a number of clusters, if the purity of the cluster is high, in other words, the sample label contained in a cluster is roughly the same, then it means high accurate in this RP by utilizing the KNN algorithm. On the contrary, if the WIFI fingerprints contained in a cluster are taken from many different RPs, which means that the purity of the cluster is low, then we named this cluster

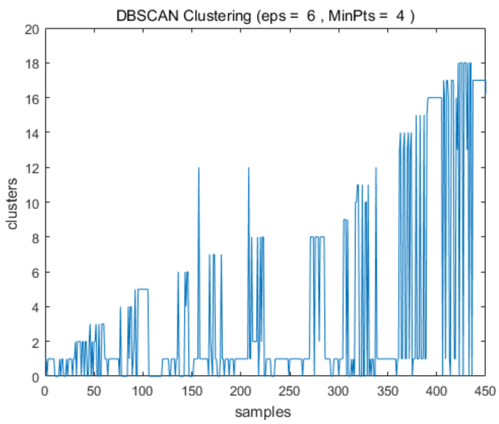


FIGURE 7. DBSCAN clustering result under the 6 eps and 4 minpts.

as insensitive regions, we should be best not utilize KNN algorithm alone to achieve indoor positioning.

The below figure 7 shows the clustering result under the 6 eps and 4 minpts, the ordinate of the figure represents the type of cluster, it is obvious that cluster 1 contains lots of samples from many different RPs.

B. KNN ALGORITHM

In the indoor positioning, the similarity of RSS fingerprints can be evaluated by calculating the Euclidean distance between them [28]. The formula is shown in below formula.

$$D\left(R, L_i\right) = \sqrt{\left(R - L_i\right)^2} \tag{1}$$

In formula 1,R is the RSS fingerprint of the estimated position, and L_i is the ith RSS fingerprint in the fingerprint library,D (R,L_i) represents the degree of similarity between R and L_i, the smaller the D (R,L_i), the more similar both two fingerprints.

When the calculation is complete, sort the D (R,L_i) from small to large and take K nearest positions, the estimated position is classified by a majority vote of its K nearest positions.

C. RANDOM FOREST ALGORITHM

When the RSS fingerprint of estimated point belongs to the insensitive cluster, the KNN algorithm is not very good to determine the position of the detection point, hence we require another method that does not rely on distance calculation to achieve indoor positioning.

Random forest algorithm is a bagging algorithm based on decision tree model, its form is represented as a set of multiple decision trees DT $= \left\{ dt_1 \ \ dt_2 \ \cdots \ dt_n \right\}$. Among them, the meta classifier dt_n is a pruned classification regression tree constructed by the CART algorithm [29], [30].

The principle is as follows.

(1) Simple random sampling I samples from the whole training set with replacement and get a new training set.

(2) Randomly extract L features from the new training set generated in step 1 and get another new training set, then train a decision tree by using this set.

(3) Repeat steps 1 and 2 until n decision trees are trained

In the final classification, each decision tree will get a classification result, utilize the majority voting method to get the final classification result of Random forest.

D. DBSCAN-KRF ALGORITHM

Integrating the above algorithms, proposing an indoor positioning algorithm based on sample space discriminant which named DBSCAN-KRF algorithm.

The steps are as follows

(1) Cluster the fingerprint library by using DBSCAN algorithm with a larger parameter, some noise samples could be detected in the clustering result, we can delete them and get a new fingerprint library.

(2) Cluster the fingerprint library obtained in Step 1 by using DBSCAN algorithm with an appropriate parameters, we can find X insensitive regions. After taking out the fingerprint samples from the insensitive regions, we can obtain X random forests Model $\left\{ R_1 \ \ R_2 \ \cdots \ R_X \right\}$ by training them with the RF algorithm.

(3) Enter the RSS fingerprint received at offline phase, use the KNN algorithm to obtain an estimated position P.

(4) Judging the first N similar fingerprint samples obtained by using the KNN algorithm. If M of them is in the sensitive region, the final estimated position is P. If M of them is in the insensitive region and most of them belongs to the *sth*, enter the RSS fingerprint into the random forest model R_S, we can obtain the final estimated position Q.

In order to describe this algorithm better, pseudo-code for DBSCAN-KRF algorithm is detailed in Algorithm 1.

V. EXPERIMENTAL RESULTS

In order to reflect the superiority of this algorithm more comprehensively, choosing KNN algorithm, random forest algorithm and WKNN algorithm to calculate their positioning accuracy under various conditions, and comparing the positioning accuracy of each algorithm. Due to the two most important factors in learning are the number of samples and the sample characteristics, and so these two factors, the number of RSS samples and the number of AP, are chosen as the condition to calculate their accuracy in the indoor positioning algorithm.

Figure 8 shows the positioning accuracy of each algorithm under different APs number. The positioning accuracy is calculated when the positioning error is less than 1.5m and the number of RSS sample in each RPs is 40. We can easily get that positioning accuracy is proportional to the number of APs in general trend. This is because the more APs there are, the more information will be contained in the positioning space. However, each AP have different effect on the indoor positioning algorithm, some APs make a great contribution to accuracy, while others make very small, even make a negative contribution.

Figure 9 shows the positioning accuracy of each algorithm under different RSS samples number. The positioning accuracy is calculated when the positioning error is less than 1.5m

Algorithm 1 DBSCAN-KRF Algorithm

INPUT: Data of online fingerprint: X_{es}
INPUT: Data of fingerprint library: $\{X_1\ X_2\ \cdots\ X_n\}$
INPUT: large DBSCAN parameters eps: ϵ_l
INPUT: large DBSCAN parameters minpts: ρ_l
INPUT: appropriate DBSCAN parameters eps: ϵ_a
INPUT: appropriate DBSCAN parameters minpts: ρ_a
INPUT: KNN parameters: K
INPUT: examine KNN result samples number: N
INPUT: threshold number: M
OUTPUT: estimated position: Q
1: set DBSCAN parameters ϵ_l and ρ_l
2: cluster $\{X_1\ X_2\ \cdots\ X_n\}$ by DBSCAN
3:delete noise samples and obtain new fingerprint library: $\{X_1\ X_2\ \cdots\ X_m\}$
4:set DBSCAN parameters ϵ_a and ρ_a
5:cluster $\{X_1\ X_2\ \cdots\ X_m\}$ by DBSCAN
6:obtain sensitive regions $\{C_1\ \cdots\ C_Y\}$and insensitive regions $\{C'_1\ \cdots\ C'_H\}$
7: while insensitive regions number $i < H$ do
8: take out fingerprint in cluster C'_i:$\{X^i_1\ X^i_2\ \cdots\ X^i_l\}$
9: training $\{X^i_1\ X^i_2\ \cdots\ X^i_l\}$ by RF algorithm
10: obtain random forests Model: R_i
11: end while
12: set KNN parameters K
13: estimated position of online fingerprint X_{es} by KNN
14: obtain estimated position: P
15: obtain similarity of fingerprint from high to low: $\{X^1\ X^2\ \cdots\ X^n\}$
16: take out the first N similar fingerprint: $\{X^1\ X^2\ \cdots\ X^N\}$
17: obtain fingerprint $\{X^1\ X^2\ \cdots\ X^N\}$ cluster: $\{C^1\ C^2\ \cdots\ C^N\}$
18: calculate and get the number of insensitive regions in $\{C^1\ C^2\ \cdots\ C^N\}$: NC
19: if NC $< M$ do
20: the final estimated position: Q:= P
21: else
22: judge the most frequently occurring insensitive cluster : sth
23: enter the X_{es} into the random forest model R_S
24: obtain the final estimated position: Q
25: end if

and the APs number is 8. We can learn from the graph that the number of RSS samples is a very important factor in the fingerprint based positioning algorithm, because with the increase of the number of samples, the accuracy of positioning algorithms has been increased, and there is no drop in accuracy. This is because the more RSS samples are collected, the more comprehensive the information of signal distribution is obtained.

It can be seen easily from the figure 8 and figure 9, DBSCAN-KRF algorithm is superior to other algorithms under the condition of different APs number or different RSS samples number.

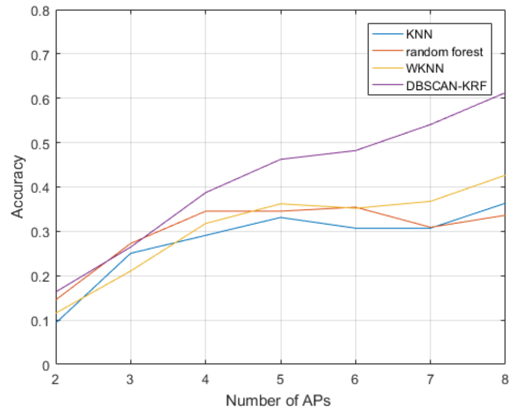


FIGURE 8. Accuracy of each algorithm under different Aps.

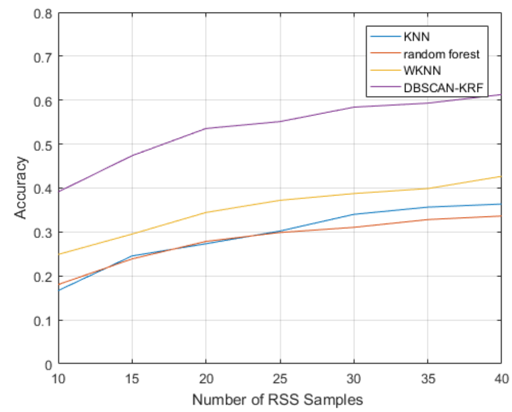


FIGURE 9. Accuracy of each algorithm under different RSS samples number.

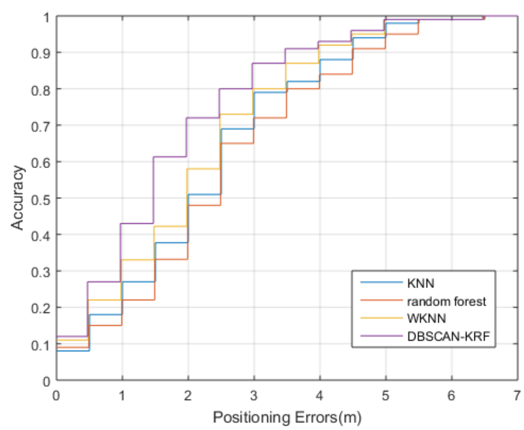


FIGURE 10. Positioning accuracy under different positioning errors.

Figure 10 is the cumulative distribution probability (CDF) of positioning errors. We can learn from the figure that the accuracy of all algorithms can reach nearly 95% when the positioning error is less than 5m, and the DBSCAN-KRF algorithm is superior to other algorithms under any positioning error.

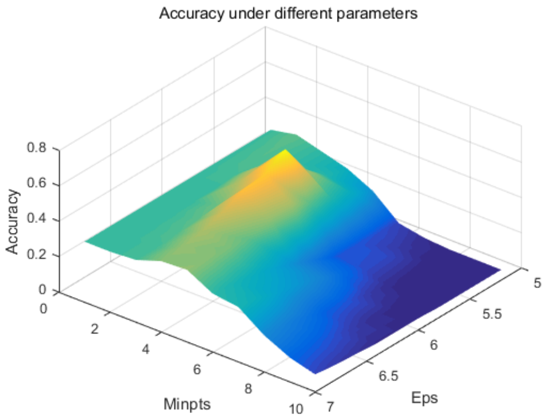


FIGURE 11. Positioning accuracy under different parameters.

Figure 11 shows the positioning accuracy under different DBSCAN parameters, when set parameters eps is 6 and minpts is 4, the highest accuracy 61.28% can be obtained. The positioning accuracy is calculated when the positioning error is less than 1.5m, the APs number is 8, and the number of RSS sample in each RPs is 40.

VI. CONCLUSION

In this study, aiming at the problem that the KNN algorithm cannot correctly judge the position information in some regions, the concept of insensitive region of RSS fingerprint signal is proposed, this region can be detected from fingerprint Library by using DBSCAN clustering algorithm. In addition, the KNN algorithm and the random forest algorithm are used to achieve the positioning in different types of regions. Experiment result shows that the proposed algorithm is superior to several other algorithms, it can improve the positioning accuracy. However, in this experiment, the area of positioning environment is not big, so this study does not take into account the positioning rapidity in large environment, moreover, the positioning target is unmoving, hence the positioning of moving targets in a large environment needs to be studied in further.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] A. Brack, "Reliable GPS + BDS RTK positioning with partial ambiguity resolution," *GPS Solutions*, vol. 21, no. 3, pp. 1083–1092, 2017.

[2] D. Zhang, J. Han, C. Li, J. Wang, and X. Li, "Detection of Co-salient objects by looking deep and wide," *Int. J. Comput. Vis.*, vol. 120, no. 2, pp. 215–232, 2016.

[3] G. Cheng, P. Zhou, and J. Han, "Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7405–7415, Dec. 2016.

[4] A. H. Ismail, R. Tasaki, H. Kitagawa, and K. Terashima, "Optimum placement of wireless access point for mobile robot positioning in an indoor environment," *J. Robot. Mechatron.*, vol. 28, no. 2, pp. 162–172, 2016.

[5] M. Aykac, E. Erçelebi, and N. B. Aldin, "ZigBee-based indoor localization system with the personal dynamic positioning method and modified particle filter estimation," *Analog Integr. Circuits Signal Process.*, vol. 92, no. 2, pp. 263–279, 2017.

[6] M. Abedian, S. K. A. Rahim, S. Danesh, and T. A. Rahman, "Compact UWB dielectric resonator antenna with WLAN band rejection," in *Proc. Int. Symp. Antennas Propag. (ISAP)*, Nov. 2015, pp. 1–4.

[7] S.-H. Fang, C.- H. Wang, T.-Y. Huang, C.-H. Yang, and Y.-S. Chen, "An enhanced ZigBee indoor positioning system with an ensemble approach," *IEEE Commun. Lett.*, vol. 16, no. 4, pp. 564–567, Apr. 2012.

[8] S. Shahidi and S. Valaee, "GIPSy: Geomagnetic indoor positioning system for smartphones," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Oct. 2015, pp. 1–7.

[9] C. Shen, Y. Li, Y. Chen, X. Guan, and R. Maxion, "Performance analysis of multi-motion sensor behavior for active smartphone authentication," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 1, pp. 48–62, Jan. 2018.

[10] O. Abdullah, I. Abdel-Qader, and B. Bazuin, "K-means-Jensen-Shannon divergence for a WLAN indoor positioning system," in *Proc. IEEE 7th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Oct. 2016, pp. 1–5.

[11] M. Jian, C.-F. Zhang, F. Yan, and M.-Z. Tang, "A global line extraction algorithm for indoor robot mapping based on noise eliminating via similar triangles rule," in *Proc. 35th Chin. Control Conf.*, Jul. 2016, pp. 6133–6138.

[12] Q. Chang, S. Van de Velde, W. P. Wang, Q. Li, H. T. Hou, and S. Heidi, "Wi-Fi fingerprint positioning updated by pedestrian dead reckoning for mobile phone indoor localization," in *Proc. China Satell. Navigat. Conf. (CSNC)*, vol. 342, 2015, pp. 729–739.

[13] G. Caso, L. de Nardis, and M.-G. di Benedetto, "A mixed approach to similarity metric selection in affinity propagation-based WiFi fingerprinting indoor positioning," *Sensors*, vol. 15, no. 11, pp. 27692–27720, Oct. 2015.

[14] L. Chen, B. Li, K. Zhao, C. Rizos, and Z. Zheng, "An improved algorithm to generate a Wi-Fi fingerprint database for indoor positioning," *Sensors*, vol. 13, no. 8, pp. 11085–11096, 2013.

[15] C. Shen, C. Liu, H. Tan, Z. Wang, D. Xu, and X. Su, "Hybrid-augmented device fingerprinting for intrusion detection in industrial control system networks," *IEEE Wireless Commun.*, vol. 25, no. 6, pp. 26–31, Dec. 2018.

[16] C. Shen, Y. Chen, X. Guan, and R. Maxion, "Pattern-growth based mining mouse-interaction behavior for an active user authentication system," *IEEE Trans. Dependable Secure Comput.*, to be published. doi: 10.1109/TDSC.2017.2771295.

[17] J. Ma, X. Li, X. Tao, and J. Lu, "Cluster filtered KNN: A WLAN-based indoor positioning scheme," in *Proc. Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Jun. 2008, pp. 1–8.

[18] Z. Wu, K. Fu, E. Jedari, S. R. Shuvra, R. Rashidzadeh, and M. Saif, "A fast and resource efficient method for indoor positioning using received signal strength," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9747–9758, Dec. 2016.

[19] T. Armstrong, A. M. Brown, P. M. Chadwick, J. A. Graham, and S. J. Nolan, "DBSCAN re-applied to pass 8 *Fermi*-LAT data above 100 GeV," *AIP Conf. Proc.*, vol. 1792, no. 1, 2017, Art. no. 070001.

[20] Y. Chen and C. Shen, "Performance analysis of smartphone-sensor behavior for human activity recognition," *IEEE Access*, vol. 5, pp. 3095–3110, 2017. doi: 10.1109/ACCESS.2017.2676168.

[21] J. Liu, C. Wu, Z. Wang, and L. Wu, "Reliable filter design for sensor networks using type-2 fuzzy framework," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1742–1752, Aug. 2017.

[22] Y. Zhao, J. Wang, F. Yan, and Y. Shen, "Adaptive sliding mode fault-tolerant control for type-2 fuzzy systems with distributed delays," *Inf. Sci.*, vol. 473, pp. 227–238, Jan. 2019.

[23] A. Abusara and M. Hassan, "Enhanced fingerprinting in WLAN-based indoor positioning using hybrid search techniques," in *Proc. Int. Conf. Commun., Signal Process., Appl. (ICCSPA)*, Feb. 2015, pp. 1–6.

[24] J. Luo and L. Fu, "A smartphone indoor localization algorithm based on WLAN location fingerprinting with feature extraction and clustering," *Sensors*, vol. 17, no. 6, p. 1339, 2017.

[25] A. Abusara, M. S. Hassan, and M. H. Ismail, "Reduced-complexity fingerprinting in WLAN-based indoor positioning," *Telecommun. Syst.*, vol. 65, no. 3, pp. 407–417, 2017.

[26] W. Chen, Q. Chang, H.-T. Hou, and W.-P. Wang, "A novel clustering and KWNN-based strategy for Wi-Fi fingerprint indoor localization," in *Proc. 4th Int. Conf. Comput. Sci. Netw. Technol. (ICCSNT)*, Dec. 2015, pp. 49–52.

[27] O. Limwattanapibool and S. Arch-Int, "Determination of the appropriate parameters for K-means clustering using selection of region clusters based on density DBSCAN (SRCD-DBSCAN)," *Expert Syst.*, vol. 34, no. 3, Jun. 2017, Art. no. e12204.

[28] X. Ge and Z. Qu, "Optimization WIFI indoor positioning KNN algorithm location-based fingerprint," in *Proc. IEEE 7th Int. Conf. Softw. Eng. Service Sci. (ICSESS)*, Aug. 2016, pp. 135–137.

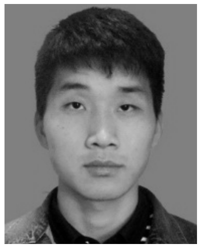
[29] S. Hariharan, S. Tirodkar, A. Porwal, A. Bhattacharya, and A. Joly, "Random forest-based prospectivity modelling of greenfield terrains using sparse deposit data: An example from the Tanami region, Western Australia," *Natural Resour. Res.*, vol. 26, no. 4, pp. 489–507, 2017.

[30] M. T. Uddin and M. M. Islam, "Extremely randomized trees for Wi-Fi fingerprint-based indoor positioning," in *Proc. 18th Int. Conf. Comput. Inf. Technol. (ICCIT)*, Dec. 2015, pp. 105–110.

KAI WANG received the B.S. degree in electronics and information engineering from Sichuan University, Chengdu, China, in 2000, and the Ph.D. degree in control theory and control engineering from Chongqing University, Chongqing, China, in 2009. He was a Visiting Scholar with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA, in 2014. He is currently an Associate Professor with the Department of Automation, Chongqing University. His current research interests include machine learning, deep learning, and the application of big data.

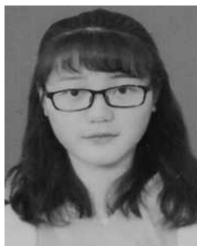
XING YU is currently pursuing the master's degree in control engineering from Chongqing University, under the Supervision of Prof. K. Wang. His research interests include machine learning and data mining.

QINGYU XIONG received the B.Sc. and M.Sc. degrees from Chongqing University, in 1986 and 1991 respectively, and the Ph.D. degree from Kyushu University, Japan, in 2002. He is currently a Professor and Supervisor for Ph.D. Student with Chongqing University. His main research interests include intelligent systems and intelligent computation, pervasive computation, and embedded systems.



QIWU ZHU is currently pursuing the master's degree with the Department of Control Engineering, University of Chongqing, under the Supervision of Prof. K. Wang. His research interests include machine learning and data mining.

WANG LU is currently pursuing the master's degree with the Department of Control Engineering, University of Chongqing, under the Supervision of Prof. K. Wang. His research interests include machine learning and data mining.



YA HUANG is currently pursuing the master's degree with the Department of Control Engineering, University of Chongqing, under the Supervision of Prof. K. Wang. Her research interests include machine learning and data mining.

LINYU ZHAO is currently pursuing the master's degree in control engineering from Chongqing University, under the Supervision of Prof. K. Wang. Her current research interests include machine learning and data mining.