Weight Matrix Analysis Algorithm for WLAN Indoor Positioning System

Lin Ma(✉), Jian Li, and Yubin Xu

Communication Research Center, Harbin Institute of Technology, Harbin, China
malin@hit.edu.cn

Abstract. Because WLAN signal strength data is vulnerable to external interference and its validity period is short-lived, it is necessary to reconstruct radio map to improve positioning performance. We add the weight data to the original fingerprint library, which is obtained by the reliability of information. Using it, we can know the importance of neighborhood points selected in the online phase and get better positioning performance. In the positioning phase, the KD tree is added to improve the positioning efficiency of the positioning algorithm. Finally, the positioning accuracy and efficiency can be improved.

Keywords: Indoor positioning · WLAN · Radio map · Weight matrix analysis

1 Introduction

With the development of positioning technology, people use positioning more and more widely. At the same time, people's demand for indoor positioning is also increasing. At present, indoor positioning technology based on WLAN [1, 2] has made great progress in the algorithm, but it still lacks specific management and construction in the use of fingerprint database. The fingerprint database usually consists of receiving signal strength information, because the fingerprint database is continuously updated. we need to continuously update signal strength information from access point (AP). These information is the uneven distribution and different quality, therefore, it is necessary to reconstruct the fingerprint database in order to reduce the uneven distribution and the poor positioning effect caused by the poor reliability information.

At the present stage, the research on the reconfiguration of Radio map is mainly aimed at improving the positioning effect and positioning efficiency of the original positioning system [3–5], several reconstruction algorithms will be introduced below. Bong [6] proposed a method to optimize the received signal strength (RSS), through smoothing the RSS processing radio map. This method can retain the positioning accuracy, simplify radio map and improve the efficiency of online positioning. Li [7] Put forward a set of good radio map reconstruction model. Removing abnormal data, dividing the region of radio map, selecting AP [8, 9] and reducing the dimension of AP in the region, using the weighted K nearest neighbor (KNN) algorithm [10] to position, positioning time, the capacity of the database are reduced when the positioning accuracy is guaranteed.

L. Meng and Y. Zhang (Eds.): MLICOM 2018, LNICST 251, pp. 317–326, 2018.
https://doi.org/10.1007/978-3-030-00557-3_32

Although the current reconstruction algorithm can make the original radio map get a good improvement in localization accuracy and computation time, but we can not guarantee the stability of the fingerprint database. When continuous abnormal signal is received over a period of time, that affects the accuracy of the positioning system, it is difficult to solve this problem easily by eliminating abnormal data. To solve this problem, a weight database is constructed, which is based on RSS similarity and AP and provide a positioning reference weight for the data in different time periods to assist on-line positioning and to improve the stability of radio map. In order to ensure the efficiency of database location, KD tree [11] is constructed based on RSS data and weight data. The following is the overall structure of this paper. In Sect. 2, we will introduce the overall model architecture of the indoor positioning system, and give the positioning algorithm and the specific function of the reconfiguration technology applied in the indoor positioning system and the composition of the reconstruction algorithm proposed in this paper. In Sect. 3, we will give the concrete algorithm of the reconstruction technology, including the establishment of the weight database and the construction of KD tree. In Sect. 4, we will implement the reconstruction technology, and compare it with the performance of the previous reconstruction technology. Finally, conclusion will be drawn.

2 Indoor Positioning System Model

Positioning based fingerprint positioning system can be divided into two phases, as shown in Fig. 1, including the offline phase and the online phase [12]. The offline phase mainly completes the establishment of radio map, and the online phase mainly completes the application of radio map to realize the matching and positioning.

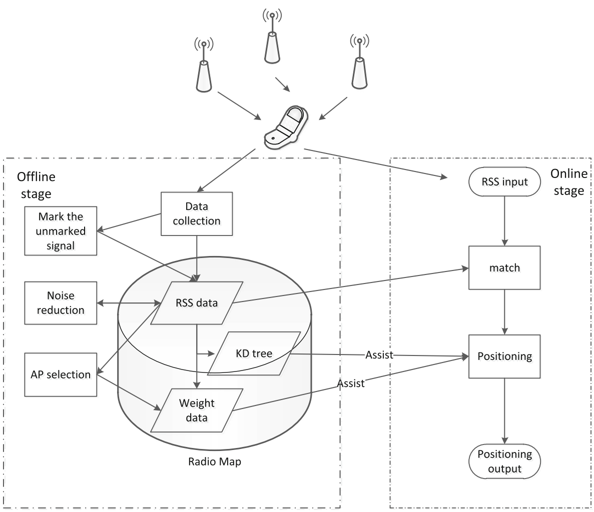


Fig. 1. Positioning system based on position fingerprint

In the offline phase, the wireless base station receives a signal from a WLAN client, and the system will establish the mapping relationship between the received signal strength of all AP in the target area and the actual physical location, and we use the mapping to construct radio map. It is assumed that there are M access points and location information for N reference points in the positioning environment. The radio map format is constructed as follows

$$RadioMap = \begin{bmatrix} L & S \end{bmatrix} \tag{1}$$

where $L \in R^{2\times N}$ is the position information and contains the reference point position coordinate, and $S \in R^{\mathrm{M}\times N}$ is the received signal strength information and contains the received signal strength of N dimension. The initial radio map is of great computational complexity and poor positioning accuracy, so it is necessary to optimize the initial Radio [13] to improve the positioning efficiency of radio map. These optimization algorithms can be classified as part of the reconstruction technology, including the noise reduction, smoothing and clustering algorithms for the fingerprint library [14]. The reconstruction technology proposed in this paper is applied to the offline phase, and the reliability analysis of RSS data is carried out during the off-line updating phase, and the weighted matrix and KD tree are constructed.

In the online phase, the positioning [15] is implemented on the basis of optimized radio map. The received data is used to position mainly by using nearest neighbor method, maximum likelihood probability method, kernel function method, neural network method and support vector regression method. In the reconstruction algorithm proposed in this paper, the weight database is involved in the assisted positioning, and the weighted KNN algorithm is used. Compared to the traditional KNN algorithm, weighted KNN increases the sample weight allocation step, but this phase takes little time comparing with the classification decision phase. The weighted KNN algorithm is applied to this subject. The redistribution of weights is realized on the objective conditions in the actual environmental factors. The reconfiguration technique proposed in this paper is based on the weight database, and the weights in the database are applied to the weighted KNN algorithm to assist the on-line positioning. Formula is as follows

$$D_k = \sqrt{\sum_{i=1}^{\mathrm{m}} \beta_i \cdot (RSS_{u,i} - RSS_{k,i})^2} \tag{2}$$

$$(x_u, y_u) = \sum_{k=1}^{n} \gamma_k \cdot \frac{1}{D_k^{\alpha}} \cdot (x_k, y_k) \tag{3}$$

where the RSS is the information in the RSS database, the β and γ are the weights in the weight database, and D is the distance between the reference points and the positioned points.

3 Reconstruction Technology

As mentioned above, in the online phase, the localization algorithm needs to create a weight database to assist the positioning. The reconstruction algorithm proposed in this paper consists of two parts, obtaining weights based on RSS and AP, and constructing KD tree. Through the above several algorithms, the original radio map is reconstructed, and the weight matrix and the KD tree are obtained. Finally, the reconstructed fingerprint database is combined with the weight matrix and the KD tree to Position online.

3.1 Weight Based on RSS

The weighting algorithm based on RSS uses the Truth Finder algorithm for reference, and obtains the reliability by comparing the similarity between RSS. Because of different mobile devices in different time periods, the reliability of accepting RSSs for each period of time is different, therefore the accuracy of the data in different time periods needs to be calculated, which can be calculated from the arithmetic mean of the probability that the data is true

$$a_w = \frac{1}{m}\sum_{k=1}^{m} p(e \le \delta) \quad w \in W \tag{4}$$

where m is the number of RSS messages obtained for each period, and the error is within the acceptable range, and δ is an acceptable error.

Obtained by the Bayesian formula

$$p(x) \propto \prod_{w \in W} \frac{a_w}{1 - a_w} \tag{5}$$

Where w is a period of time, W is all time periods. Let

$$c(x) = \ln \prod_{w \in W} \frac{a_w}{1 - a_w} = \sum_{w \in W} \ln \frac{a_w}{1 - a_w} \tag{6}$$

$$q = \ln \frac{a}{1 - a} \tag{7}$$

Based on (6) and (7), we can have:

$$c(x) = \sum_{w \in W} q_w \tag{8}$$

The normalized probability distribution is obtained, which is added to the weight matrix as a weight γ.

$$p(X{=}x) = \frac{c(x)}{\sum c(y)} \tag{9}$$

where $c(x)$ can be corrected by the similarity between RSSs, the correction process is as follows.

Signal similarity is

$$r(x,x') = \frac{xx'}{|x||x'|} \tag{10}$$

The corrected credibility is

$$c'(x) = c(x) + \sum_{e(x)=e(x')} (r(x,x') - \rho)c(x') \tag{11}$$

where ρ is the similarity parameter, and by adjusting the parameter, we can get the effective weight. This parameter is selected according to the positioning environment. Eventually, the weight γ can be obtained from the posterior probability.

3.2 Weight Based on AP

Taking account of the positioning environment, the performance of each AP is different. In the positioning process, the poor performance of the AP will often cause the positioning accuracy decreased. So each AP needs a weight to express its performance is good or bad in a particular environment, and we use AP selection algorithm to determine the weight.

In the AP selection algorithm, the simplest algorithm is to keep all APs as samples, but this does not play a role in improving the system. The MaxMean algorithm was proposed by Youssef. We need to average the RSS values of the signals received by all APs, and then select several APs with large RSS meanings as optional access points for that location and provide weight to these APs.

Another AP selection algorithm is the InfoGain algorithm, which is based on the gain of information entropy, which is presented by Chen of Hong Kong University of Science and Technology. Firstly, the stochastic entropy of radio map is calculated, and then the conditional entropy of different APs is calculated. The AP is sorted by calculating the difference between stochastic entropy and conditional entropy. The AP with high information entropy gain is used to construct the fingerprint library.

$$IG(AP_i) = H(G) - H(G|AP_j) \tag{12}$$

among them:

$$H(G) = -\sum\nolimits_{g=1}^{n_i} p(G_g)\log_2 p(G_g) \tag{13}$$

$$H(G|AP_j) = -\sum\nolimits_{v}\sum\nolimits_{g=1}^{n_i} p(G_g, AP_j = v)\log_2 p(G_g|AP_j = v) \tag{14}$$

where $H(G)$ is the gain of the information entropy of the AP, $IG(AP_i)$ is the information entropy when there is no AP information, $p(G)$ is the information entropy for the AP information, and $p(G)$ is the probability of judging the various positions. The weight is obtained from the gain of the information entropy.

There is also a method based on the stability of the AP, firstly we use the RSS variance information to determine an amount that represents the AP stability. At the same time, considering the relationship between the intensity of the signal and the frequency of the occurrence of the signal, the energy information of the signal is reflected by a frequency. Combine two aspects to obtain the value of the AP stability as a weight. The formula is as follows

$$Sta(AP_i) = \frac{1}{\frac{1}{N}\sum\limits_{j=1}^{N}(RSS_j - \overline{RSS})^2} \cdot \frac{N_i}{Sum_{k=1}^{n}(N_k)} \tag{15}$$

where N_i is the frequency at which each AP appears.

Using the above algorithms, we can get the weight of each AP in the region, which is the weight β mentioned above.

The weighted KNN algorithm described above requires the traversal of the entire database when we need to obtain the nearest neighbor reference point, so it takes a lot of time in the online positioning phase. The weighted KNN algorithm is combined with the KD tree and has a very high efficiency. By creating the KD tree, it can save the time of calculation.

KD tree is a high-dimensional index tree data structure, commonly used in large-scale high-dimensional data space for the nearest neighbor search and approximate nearest neighbor search, such as high-dimensional image feature vector for K proximity searching and matching in image retrieval and recognition. When a K-nearest neighbor of a reference point is required, it is only necessary to trace along the branches of the KD tree to find the nearest few data.

KD tree construction algorithm: select the dimension with the largest variance in the k-dimensional data set, and then select the median m in this dimension as pivot to divide the data into two sub-sets, while creating a tree node for storage. This step is repeated until all sub-sets can not be divided. If a subset can no longer be divided, the data in the sub-set is saved to the leaf node.

4 Implementation and Performance Analysis

4.1 Experiment Environment

The experimental environment is located at 12th floor, 2A Building, Science and Technology Park, Harbin Institute of Technology. The height of this floor is 3 m, the area is 66.43 × 24.9 m^2, with 19 laboratories and a conference room and a table tennis room. The experimental facilities were spread across the 12th floor laboratory and were fixed at a height of 2 m. The device supports IEEE 802.11 g standard, and its transmission rate is 54 Mbps. The receiver is 1.2 m from the ground.

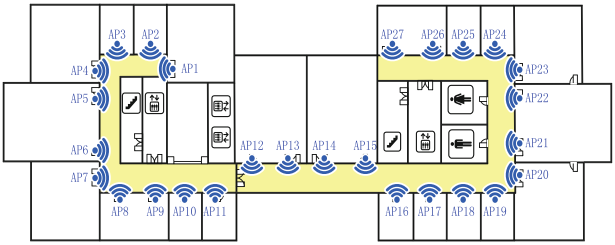


Fig. 2. Experiment environment

In order to test the universality of the reconstructed model, we select the 12th floor corridor as the experimental area, so that we can get more signals from the AP in the experiment, and the AP selection algorithm can be effectively tested. 27 APs evenly distribute in the laboratories, as shown in Fig. 2, to ensure that any location in the 12 layer can have more AP. Use a laptop computer as a data collector to collect fingerprint library data and test data.

In the laboratory corridor, the yellow area in the figure, evenly select the reference point within every 0.5 m, and select 823 reference points and rely on these reference points to build radio map. Firstly collect 420 data at each sampling point, and then remove the 20 data in the first and tail, and finally use the remaining 400 data to build the initial radio map to ensure that the initial radio map has a better positioning effect.

In addition to building a radio map, we also need to get the test data used in the online positioning phase. We collect the data in a random location, collected in five time periods. We collect data in five time periods, because the need for different time periods to grant different weights to the original fingerprint library and updating operation in the establishment of the RSS weight. Assume that the test data selected in the five time periods are not relevant and are randomly distributed in the positioning area. We need to collect 200 times to get the final test data in each time period, and collect data in a random location.

4.2 Performance Analysis

Based on different test data for several experiments, Fig. 3 is the cumulative error probability obtained three reconstruction algorithms. The reconstruction algorithm based on AP selection is the reconstruction algorithm mentioned above. Using the noise reduction algorithm and AP selection algorithm to optimize the original radio map, we can see that it can improve the positioning accuracy after removing the abnormal information and the selection of feasible AP, cumulative probability within 2 m increased by 8%. The reconstruction algorithm based on weight matrix proposed in this paper can more effectively distinguish the data confidence in each time. Compared with the noise reduction algorithm, the positioning accuracy is improved, and the cumulative probability of positioning within two meters is increased by 3%.

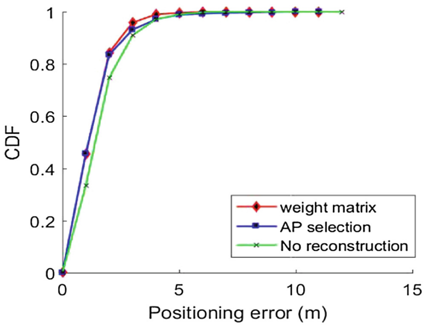
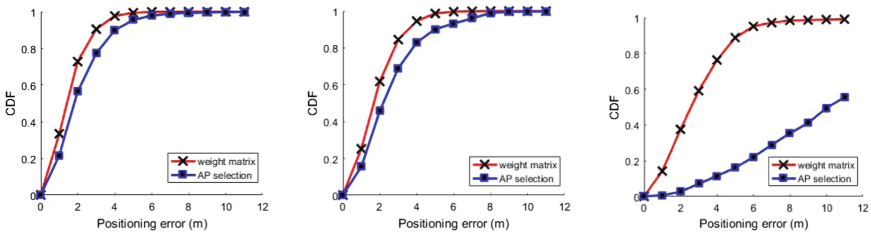


Fig. 3. Different reconstruction algorithm

Next, for different noise powers, the performance of the reconstructed technique is compared. When the power spectral density is gradually increased and the system parameters are kept as the optimal parameters, the experimental comparison is carried out. Figure 4 shows the cumulative error probability of the positioning error in the environment where the noise power spectral density is increased by 3, 4 and 10. It can be seen from the trend of curve with the noise enhancement, the maximum positioning error is increasing. In the low-noise environment, the maximum positioning error of the reconstruction algorithm based on weight matrix is better than that of AP-based



(a) Increase 3db (b) Increase 5db (c) Increase 10db

Fig. 4. Different noise environments

reconstruction algorithm, and the positioning error is reduced by 3.5 m. It can be seen that the reconstruction algorithm based on AP can not guarantee the lower positioning error and can not meet the requirement of location, when the noise power spectral density is increased by 10. However, the reconstruction algorithm based on weight matrix has better positioning effect.

The positioning efficiency of the two reconstruction algorithms is compared below. Table 1 is the comparison of the positioning usage time under different fingerprint libraries. The first reconstruction scheme uses zoning to position, using multi-step positioning, which can effectively improve the positioning efficiency. The second location is the KD tree proposed in this paper. As can be seen from the experimental results, the use of KD tree can save a lot of time in the positioning phase, and with the increasing number of fingerprint libraries, the time saved is also rising compared to the former reconstruction algorithm.

Table 1. KD tree experiment

Fingerprint library size (number of RSS)	1000	2000	5000	10000	20000
Using zoning (s)	0.472	0.939	6.976	14.558	30.320
Using the KD tree (s)	0.145	0.208	0.283	0.510	1.068

5 Conclusion

In this paper, a new method is proposed to solve the problem that the stability of the positioning system is degraded by the long time abnormality signal of radio map. The original fingerprint library is reconstructed by constructing the weight matrix and KD tree of radio map. And we compare the performance of the reconstruction algorithm in the laboratory environment. The accumulated positioning error within two meters is increased by 11% and increased by 3% to comparing the reconstruction algorithms based on AP selection. The algorithm presented in this paper uses the weight matrix provided by RSS reliability analysis to assist in positioning, and the performance of resisting noise has been improved. In view of the positioning efficiency, the proposed positioning algorithm combined with the KD tree simplifies the complexity of the nearest neighbor search algorithm. It can be seen from the experimental results that the positioning time is shortened.

Acknowledgment. This paper is supported by National Natural Science Foundation of China (61571162), Ministry of Education - China Mobile Research Foundation (MCM20170106) and Heilongjiang Province Natural Science Foundation (F2016019).

References

1. Gartner, G., Ortag, F.: Advances in location-based services. Lect. Notes Geoinformation Cartogr. **5**(8), 97–106 (2014)
2. Feng, C., Au, W.S.A., Valaee, S., et al.: Received-signal-strength-based indoor positioning using compressive sensing. IEEE Trans. Mob. Comput. **11**(12), 1983–1993 (2012)
3. Baala, O., Zheng, Y., Caminada, A.: The impact of AP placement in WLAN-based indoor positioning system. In: Eighth International Conference on Networks, pp. 12–17. IEEE Computer Society (2009)
4. Pan, J.J., Pan, S.J., Yin, J., et al.: Tracking mobile users in wireless networks via semi-supervised colocalization. IEEE Trans. Pattern Anal. Mach. Intell. **34**(3), 587–600 (2012)
5. Au, A.W.S., Feng, C., Valaee, S., et al.: Indoor tracking and navigation using received signal strength and compressive sensing on a mobile device. IEEE Trans. Mob. Comput. **12**(10), 2050–2062 (2013)
6. Bong, W., Kim, Y.C.: Reconstruction of radio map from sparse RSS data by discontinuity preserving smoothing. In: ACM Research in Applied Computation Symposium, pp. 227–231. ACM (2012)
7. Li, X., Deng, Z.: Radio map reconstruction technology in indoor fingerprint localization algorithm (2012)
8. Deng, Z., Ma, L., Xu, Y.: Intelligent AP selection for indoor positioning in wireless local area network. In: International ICST Conference on Communications and Networking in China, pp. 257–261. IEEE Computer Society (2011)
9. Umair, M.Y., Xiao, D., Li, A., et al.: Access point selection for indoor positioning in a WLAN environment using an algorithm based on RSSI and dilution of precision. Environ. Entomol. **26**(3), 91–99 (2014)
10. Yang, L., Chen, H., Cui, Q., et al.: Probabilistic-KNN: a novel algorithm for passive indoor-localization scenario. In: Vehicular Technology Conference, pp. 1–5. IEEE (2015)
11. Chen, X.K., Liu, Z.S.: K nearest neighbor query based on improved Kd-tree construction algorithm. J. Guangdong Univ. Technol. **31**, 119–123 (2014)
12. Kaemarungsi, K., Krishnamurthy, P.: Analysis of WLAN's received signal strength indication for indoor location fingerprinting. Elsevier Science Publishers B. V. (2012)
13. Lee, M., Han, D.: Voronoi tessellation based interpolation method for Wi-Fi radio map construction. IEEE Commun. Lett. **16**(3), 404–407 (2012)
14. Abubaker, M., Ashour, W.: Efficient data clustering algorithms: improvements over kmeans. Int. J. Intell. Syst. Appl. **5**(3), 37–49 (2013)
15. Wang, L., Wong, W.C.: A RSS based statistical localization algorithm in WLAN. In: International Conference on Signal Processing and Communication Systems, pp. 1–5 (2012)