WKNN indoor location clustering algorithm with triangle correction

Di Zhou[1, a]**,** Haiyan Lan[1,b] **,** Guoyin Zhang[1,c] **,** Xuesong Ma[2,d]**,** and Eryue Liang[1,e]

[1] College of computer science and technology, Harbin Engineering University, Harbin, Heilongjiang, 150001, China.

[2] Heilongjiang Academy of Sciences, Harbin , Heilongjiang 150001, China

[a]409310423@qq.com , [b] lanhaiyan@hrbeu.edu.cn , [c] zhangguoyin@hrbeu.edu.cn, [d]404361372@qq.com, [e]ley95@qq.com

Keywords: indoor location, location algorithm, triangle correction

Abstract. Aiming at the problem of low precision in the positioning stage of fingerprint location positioning technology based on WiFi, through the comparative study of the nearest neighbor classification algorithm (NN),K nearest neighbor classification algorithm (KNN) and Bias algorithm, and then we present a kind of weighted KNN algorithm of based on triangle correction. On the one hand, the algorithm uses the difference between the AP signal intensity value of the node and the fingerprint node to be taken as the weighting factor, and the positioning accuracy of KNN is improved by the contribution ratio of different fingerprint nodes; On the other hand, we further improve the positioning accuracy by selecting three nearest neighbor points which the unknown node must be in the triangle composed of these three points. Finally, the simulation results showed the effectiveness of the algorithm.

Introduction

Indoor positioning technology has been getting more and more attention in recent years. Taking into account the cost, operability, universality and a series of needs, positioning technology based on WiFi location fingerprint has been recognized by a wide range of scholars. As a rusult, the accuracy of fingerprint positioning will be more and more attention. WiFi-based location fingerprint positioning technology is divided into two stages[1]: offline fingerprint library construction and online positioning. In the first phase, the fingerprint database is established by collecting the signal strength from AP nodes received by the collection point. In the second stage, by matching the signal strength value of the AP nodes with the fingerprint library, the coordinate data of the nodes to be located can be estimated. The more dense the fingerprint nodes arrange, the higher the positioning accuracy is[2]. In the case of the fingerprint library, the positioning algorithm of the online positioning phase directly affects the positioning accuracy. At present, the typical localization algorithms are nearest neighbor classification algorithm (NN), K nearest neighbor classification algorithm (KNN), Bayesian algorithm and neural network algorithm. On the basis of these classic algorithms, many scholars have innovated a lot. An indoor location system based on multilayer artificial neural network (ANN) with area division is proposed by Zhou,it gets a great effect[3]. CY Tsai presents a novel location determination mechanism that uses an indoor WLAN and back-propagation neural network (BPN) that Improves the estimation accuracy[4]. WLAN hybrid indoor location method based on fuzzy c-mean clustering (FCM) and artificial neural network (ANN) is proposed by Y Xu, bridges the gap of the generalization ability of the feed-forward ANN with three layers [5]. Based on the research and analysis of the existing localization algorithm, this paper proposes a triangular correction weighted KNN algorithm.

Fingerprint library construction

Considering the universality of the construction of the fingerprint library, and the focus of this paper is location algorithm, we choose to use the attenuation factor model to build the fingerprint library. First, define the number, spacing and position of the fingerprint nodes on the simulation site, and then

calculate the distance between the AP nodes and the fingerprint nodes. Using attenuation factor model to calculate the intensity loss of the signal transmitted by the AP node to each fingerprint node. Finally, the signal strength receives by fingerprint node from each AP node can be obtained. These signal strength and the AP node ID together constitute a fingerprint library.
The attenuation factor model is:

$$P_L(d) = P_L(d_0) + 10\lambda \log(\frac{d}{d_0}) + FAF \tag{1}$$

λ represents the single-layer path loss factor, FAF represents the loss of penetrating the building.This paper select the third floor of the 21a teaching building of Harbin Engineering University as the experimental site. The simulation map shown in Fig. 1.
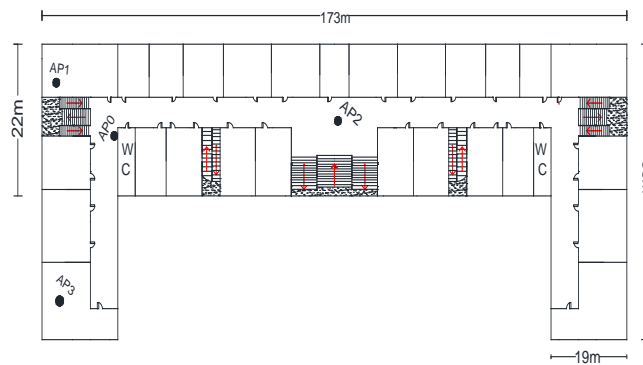

Fig. 1 21a teaching building third floor simulation map

Typical positioning algorithm

K Nearest Neighbor Algorithm.K Nearest Neighbor algorithm is the reinforcement of Nearest Neighbor (NN) Algorithm, and the difference between NN algorithm and KNN algorithm is to select K nodes and fingerprint nodes which are closest to the positioning node in the fingerprint library, and then use the centroid algorithm or the mean method to calculate the coordinates of the estimated node. In order to further improve the accuracy of the algorithm, a weighted K-nearest neighbor classification algorithm (WKNN) is proposed.
Naive Bayesian algorithm. Bayesian algorithm is a probabilistic algorithm, in the WiFi-based fingerprint positioning, the main use of the Bayesian formula is the transformation of it.

$$P(B|A) = P(A|B) * P(B) / P(A) \tag{2}$$

The goal of the algorithm is to find a position l under the condition that the signal vector $r(\langle WF1, RSSI1 \rangle \langle WF2, RSSI \rangle ... \langle WFN, RSSIN \rangle)$ transmitted by each AP node is obtained so that the probability that the signal vector near r reaches the maximum, then get max(P(r|l)). After Bayesian conversion:

$$\max(P(l|r)) = \max\left(\frac{P(r|l) * P(l)}{P(r)}\right) \tag{3}$$

In the formula, P(r) is the probability of occurrence of the signal vector, and for each measurement, P(r) is a constant that can be omitted and the probability of occurrence of the position l is assumed to be equal. So only need to calculate max(P(r|l)), which means find a point l, making the probability of occurrence of signal r the largest.

Optimized WKNN algorithm

Triangle correction algorithm.KNN algorithm is to retrieve the fingerprint library, select K points in the fingerprint library that are closest to the signal strength received by the positioning point, and then take the mean of the coordinates of the K points (in this article, K = 3). The ideal situation with the KNN algorithm is to match the three points evenly distributed around the anchor point, as shown in Fig. 2.

In the figure, the red cross symbol is the positioning point and the black dot is the matching three fingerprint nodes. It can be seen from the figure that the coordinates of the three fingerprint nodes are close to the positioning point. However, due to the complexity of the positioning environment is different, and the number and location of the AP node configuration is also different, so it can not guarantee that the matching of the three nodes evenly distributed around the positioning point, it is difficult to exclude the impact of singularity, as shown in Fig. 3, Fig. 4.
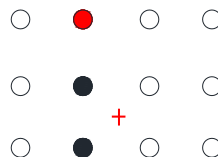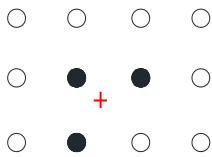

Fig. 2 KNN Match Fingerprint Ideal Distribution   **Fig.** 3 KNN Matching Fingerprint Singular Point Example I   **Fig.** 4 KNN Matching Fingerprint Singular Point Example II

The red dots in the figure are the singular fingerprint nodes. Because the interference caused by the singularity makes the obtained coordinates mean and the actual positioning point has a great deviation. In order to solve the problem of singular point interference in KNN localization, this paper proposes a method of trigonometric correction. Replace traditional KNN algorithm (matches three closest fingerprint nodes) with matching three fingerprint nodes, which must be ensured that the positioning point receives the signal strength at the closest point in the triangle formed by the three-point connection). The specific algorithm is as follows:

I. Sort the nodes of the fingerprint library according to the difference between the AP nodes' signal strength.

II. To determine whether the first three nodes meet the conditions (this article uses the slope method to determine whether the positioning point in the triangle, according to the three nodes of the y-axis coordinate size is divided into six cases).

III. Move the third node back, repeat step II until you find the three nodes that meet the criteria.

Weight selection.WKNN is an improved algorithm for KNN to further improve accuracy. In this paper, the weight is selected as the difference between the signal strength corresponding to each fingerprint node and the sum of the signal strength of each AP node.

Algorithm simulation comparison

In this chapter, the algorithms of NN, KNN, Bayes, WKNN and WKNN based on triangle correction are simulated and compared. The simulation scene is half of the three floor of the 21a teaching building of Harbin Engineering University.Randomly generate a point in the map as a true location to be positioned. The results of various algorithms are shown in Fig5.

Intuitively, the optimized KNN algorithm is closer to the actual positioning point, and the positioning accuracy is higher. And the positioning accuracy of the KNN is worse than NN,that confirms the error of KNN singular point interference,and in the case of Same singular point error , WKNN still has higher accuracy than KNN algorithm.Select 58 groups of positioning points, using the optimization algorithm and the KNN algorithm for visual contrast, as shown in Fig. 6.It can be seen from the figure that the optimization algorithm proposed in this paper has higher fitting degree than KNN, and the positioning accuracy is higher.Two hundred random positioning points are selected, and the location of each algorithm is estimated by using the five algorithms respectively. The error of each algorithm is displayed in the form of probability according to the interval, as shown in Fig. 7.

It can be seen that the accuracy of the NN algorithm is the lowest, the error in the range of less than two meters of the two hundred groups of data using the NN algorithm to locate is about 73%. KNN algorithm compared to NN algorithm has been improved, error within two meters to reach the point of 80%,and the point of high accuracy within the interval of two meters is more than NN. The accuracy of the Bayes algorithm is between KNN and WKNN algorithm which is used in this paper, Error within two meters of the basic point can reach up to 90%, the accuracy is obviously improved. Using this optimization algorithm to locate the location of the image can be seen from the map has greatly improved the accuracy. In the two hundred sets of positioning points, the error in the range of 0.5 meters to reach more than 60%, the error in the range of 0-0.5 m is also much more than other algorithms. Therefore, the optimization of the positioning accuracy is confirmed from two aspects. The average positioning error of the five algorithms is as follows:nn=1.579783, knn3=1.460857,wknn3=1.122169,bayes=1.226688,youhua=0.814404.In the case of the fingerprint node spacing of 3 meters, the average error of the algorithm based on triangle correction can reach 0.81M, which is much better than other algorithms.
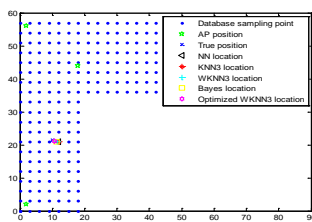
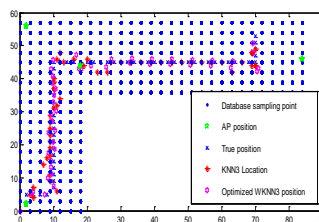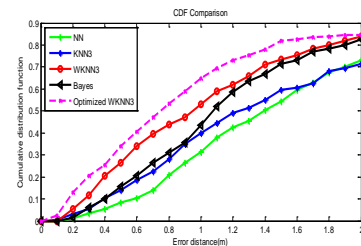|		
Fig. 5 positioning results of various algorithms	**Fig.** 6 Comparison of optimization algorithm and KNN	**Fig.** 7 error curves of the five algorithms CDF

Summary

In this paper, a new WKNN positioning algorithm based on triangle correction is proposed，based on the comparison of the typical localization algorithms.One side, in order to get rid of the singular point resulted in the different environment and the different configuration of AP,the method of triangle correction is proposed. Taking three nearest neighbors points as an example,it requires the point will be located must be in the triangle area which constructed by the three neighbors points.The other side, the difference between the signal intensity of each fingerprint node and the signal intensity of each AP node is weighted as the weight to provide different coordinate contribution. Finally, through the simulation comparison with NN algorithm, KNN algorithm ,Bayes algorithm and WKNN algorithm,the optimization of the triangle correction WKNN algorithm is confirmed, the positioning accuracy is greatly improved.When the space of the fingerprint nodes is three meters, The positioning error in the range of 2M is more than 90%, in the range of 0.5M is more than 60%.

Acknowledgements

References

[1] Karimi H. Advanced location-based technologies and services[J]. Taylor & Francis, 2013.

[2] Y.T.Chan and K.C.Ho.A Simple and Efficient Estimator for Hyperbolic Location.IEEE Trans.On Signal Processing,1994,42(8):1905-1915P

[3] Zhou, Yubin, Tang. Multilayer ANN indoor location system with area division in WLAN environment[J]. Journal of Systems Engineering and Electronics, 2010, 21(5):914-926.

[4] Tsai C Y, Chou S Y, Lin S W, et al. Location determination of mobile devices for an indoor WLAN application using a neural network[J]. Knowledge and Information Systems, 2009, 20(1):81-93.

[5] Xu Y, Mu Z, Lin M. Hybrid FCM/ANN indoor location method in WLAN environment[C]// Information, Computing and Telecommunication, 2009. YC-ICT '09. IEEE Youth Conference on. IEEE, 2009:475-478.