Selective neural network in training set on indoor positioning algorithm

Jin Ren¹, ², a), Guohao Ma¹, and Libin Ji¹

¹ School of Information Science and Technology, North China University of Technology, Beijing 100144, China
² Beijing Key Laboratory on Integration and Analysis of Large-scale Stream Data, Beijing 100144, China
a) rj@ncut.edu.cn

Abstract: We propose a new method to combine the Knn algorithm with the neural network. Due to the size of the training set affects the accuracy of the training results, we propose a neural network with selective training set. Its main method is changed the size of the training set by a control variable to get the most suitable training data. The advantage of this method is that the training set can be adjusted to achieve a trade-off between the positioning accuracy and the training set size. The result of Knn algorithm positioning is more accurate by adding BP (Back Propagation) selective neural network algorithm.

Keywords: indoor positioning, Knn, wifi, neural network

Classification: Wireless Communication Technologies

References

[1] Y. Teramoto and A. Asahara, “Wireless LAN based indoor positioning using radio-signal strength distribution modeling,” Proc. 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1–7 (2012). DOI: 10.1109/ipin.2012.6418890
[2] Z. Wu, R. Cai, S. Xu, and X. Wu, et al., “Research and improvement of wifi positioning based on k nearest neighbor method,” Computer Engineering, vol. 3, pp. 289–293 (2017).
[3] X. Ge and Z. Qu, “Optimization WIFI indoor positioning KNN algorithm location-based fingerprint,” Proc. 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), pp. 135–137 (2016). DOI: 10.1109/icsess.2016.7883033
[4] W. Xue, X. Hua, Q. Li, W. Qiu, and X. Peng, “Improved clustering algorithm of neighboring reference points based on KNN for indoor localization,” Proc. 2018 Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS), pp. 1–4 (2018). DOI: 10.1109/upinlbs.2018.8559874.
[5] M.B. Afousi and M.R. Zoghi, “Wi-Fi RSS indoor positioning system using online layer clustering and weighted DCP-KNN,” Proc. 2018 Iranian Conference on Electrical Engineering (ICEE), pp. 710–715 (2018). DOI: 10.1109/icsee.2018.8472473
[6] T. Liu, H. Mei, Q. Sun, and H. Zhou, “Application of neural network in fault location of optical transport network,” China Communications, pp. 214–225
1 Introduction

With the development of modern technology, wifi has become a necessary facility for all occasions. In order to provide wifi location services, APs (access Points) are generally placed indoor for communication purposes. In some cases, those access points are used for positioning services. The wifi terminal can continuously scan and collect surrounding AP signals and obtain its signal strength, collect AP signals through the mobile terminal and measure the location by calculation. Traditional time and angle based positioning methods such as AOA (Angle of Arrival), TDOA (Time Difference of Arrival) are not suitable for wifi [1]. Fingerprint-based positioning methods are widely studied and adopted in the location of indoor wifi scenes [2].

2 Traditional algorithm

2.1 Knn algorithm

Knn fingerprint positioning algorithm is the most commonly used algorithms in wifi positioning [3]. The algorithm is shown in Eqs. (1) and (2).

\[ d_j = \left( \sum_{i=1}^{n} (r_{ss_i} - \overline{r_{ss_i}})^2 \right)^{\frac{1}{2}} \]  

(1)

\[ (\bar{x}, \bar{y}) = \frac{1}{k} \sum_{i=1}^{k} (x_i, y_i) \]  

(2)

where \( n \) denotes the count of access points detected on the targeted location, \( r_{ss_i} \) represents the signal strength of the test point which received by APs. \( \overline{r_{ss_i}} \) represents the signal strength of the reference point’s fingerprint in the fingerprint database. \( d_j \) represents the Euclidean distance between the test point and the \( j \)-th reference point. \( k \) represents the count of reference points, \((\bar{x}, \bar{y})\) represents the location coordinate of the reference point \( i \). We found through research that many scholars have proposed the method to improve indoor fingerprint location. Weixing Xue utilized the k-means clustering algorithm to analysis the geometric proximity between RP and test point in the online phase to improve the performance of KNN algorithm [4]. Mohsen proposed a Weighted Differential Coordinate Probabilistic-KNN (WDCP-KNN) method based on probabilistic weighting of generalized Reference Points and differential coordinates [5].

2.2 BP neural network

BP neural network is currently the most widely used network. It consists of the input layer, the hidden layer, and the output layer [6]. The relationship between the input
layer and the output layer neurons is in Eq. (3).

\[ y = \sum_{i=1}^{n} w_i \cdot x_i + b \]  

(3)

where \( x \) is the input, \( y \) is the output, \( b \) is the bias matrix, and \( w \) is the weight of each input layer. The BP neural network also needs to be added on the activation function. Adding an activation function can add nonlinear factors and solve problems that cannot be solved by linear models [7]. BP neural network can automatically extract the rules between input data and output data through learning, and store the learning content in the weight of the network, which has strong adaptive ability.

3 Proposed method

In order to get the training set data better, we propose a new neural network algorithm using the selective neural network. The condition for judging the fingerprint point in the circle is in Eq. (4).

\[ r < \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \]  

(4)

where \( x_i, y_i (i = 1, 2, \ldots, n) \) represent the coordinates of the \( i \)-th fingerprint point. \( x_0, y_0 \) represent the coordinates calculated by Knn algorithm. \( r \) is the set radius. The size of \( r \) depends on the size of the indoor environment. Calculate the distance between all the fingerprint points in the database and the anchor points. Select all the fingerprint points whose distance is less than \( r \) as the training set of BP neural network. The size of \( r \) controls the amount of data in the training set.

The size of \( r \) affects the results of BP neural network training. So we need to find the suitable \( r \) by comparison. The method adopted is set the value of \( r \) differently based on the distance between neighbor fingerprint points. Take the distance of 1 times, 2 times, 4 times, 6 times, and the best \( r \) is obtained by comparison.

4 Experiment

In the BP neural network, there are 5 hidden layers, the neurons in the input and output layers are 6 and 2. The training function select the gradient descent function. First, we analyze the positioning effect of the improved Knn algorithm and other positioning algorithms. Select \( k=5 \) by control variable in Knn algorithm. We choose \( r=2m \) in the selective neural network training set. The distance error probability distribution of the proposed Knn positioning algorithm and other algorithms are shown in Fig. 1. The RMSE (Root Mean Squared Error) of the improved Knn algorithm is 2.64 meters, the Rmse of the traditional Knn algorithm is 2.89 meter. From the Fig. 5 we can see that the positioning performance of the improved Knn algorithm is better than traditional Knn algorithm, improved clustering algorithm [4] and WDCP algorithm [5]. The improved algorithm’s RMSE is 0.25m lower than traditional Knn algorithm. 70% of test points distance error are within 3 meters.

Compared with the improved clustering algorithm and the WDCP algorithm, the probability distribution of the improved Knn algorithm with a positioning error within 1 meter is higher, and only 30% of the test points have distance error more
than 3 meters. This is mainly because the improvement of the two algorithms is mainly in the Knn algorithm, but the neural network improvement algorithm can take into account the environmental factors near the positioning point. This is the advantage of neural network algorithms.

Then we adjust the value of $r$ and analyze the positioning results. We choose $r=1, 2, 4, 6$ (m). The average amount of data in their circles is 4, 12, 50, 98. The error distribution of the four cases is shown in Fig. 2. When $r=2$m, the RMSE of the positioning coordinates and the original coordinates is 2.64m. Among the positioning results of all points, 50% of the points distance error are within 2 meters, and 70% of the points distance error are within 3 meters. Through experiments, we found that if $r$ is too small, the training set will be incomplete. If $r$ is too large, there will be too much data in the training set, and unnecessary data will affect the training results. So getting the suitable $r$ is an important step to improve the positioning
performance. This is the advantage of selective neural networks training set. From Fig. 3, we can see that as $r$ increases, the convergence value of the neural network is smaller. The magnitude of change is gradually reduced, and finally leveled off. This means that the more training data, the smaller the convergence value, and it is not directly related to positioning accuracy.

![Fig. 3. Comparisons of convergence values under different values of $r$](image)

5 Conclusion

The improved Knn positioning algorithm uses selective BP neural network training set to get a more accurate position. By changing the size of $r$ to control the number of training sets, the improved Knn algorithm is more flexible. It only need a small amount of training set to achieve the best training results. In this test, the accuracy of the improved Knn algorithm at $r=2$ is 0.25m higher than other algorithms.

Acknowledgments

This work was supported in part by 2020 Science and Technology Activities Project for College Students of North China University of Technology under Grant 218051360020XN114/007, in part by 2020 Beijing College Students Innovation and Entrepreneurship Training Program Project under Grant 218051360020XN214, in part by Beijing Urban Governance Research Center Project under Grant 20XN241, and in part by Beijing Outstanding Young Backbone Individual Project under Grant 401053712002.