American Journal of Information Science and Computer Engineering
Vol. 7, No. 1, 2021, pp. 1-5
http://www.aiscience.org/journal/ajisce
ISSN: 2381-7488 (Print); ISSN: 2381-7496 (Online)

Indoor Multi-floor Localization Based on Bluetooth Fingerprint and LDA

Shi Chen[1, 2, *]

[1]School of Computer Science and Technology, Nanjing University of Technology, Nanjing, China
[2]School of Information Engineering, Yancheng Teachers University, Yancheng, China

Abstract

With the rapid development of the information age, people are no longer satisfied with outdoor location-based services, and indoor localization has become a hot topic of discussion. Related technical personnel have also actively explored indoor positioning technology. At present, many indoor localization scenes are no longer a single single-layer environment, and position estimation is also required in a multi-layer environment. However, the existing work shows certain limitations to the problem of floor localization accuracy and computational complexity. This paper proposes a localization system capable of floor recognition. The system is divided into offline phase and online phase. In the offline phase, we deploy the Bluetooth APs to collect RSS signal strength values at the planned collection points and establish a local fingerprint database. Then use linear discriminant analysis to establish a floor recognition model. In the online stage, the floor location is determined first, and then the specific location of the target is obtained through the improved KNN algorithm. We collected real experimental data on two floors. Experimental results show that we can quickly and accurately locate a floor through a small amount of AP fingerprint information, reduce the complexity of localization calculations, and accurately locate the target's specific location on the floor.

Keywords

Indoor Localization, Multi-floor Recognition, Wireless Network, Floor Classification

Received: December 11, 2020 / Accepted: January 16, 2021 / Published online: March 17, 2021

1. Introduction

Most of the life of modern urban people is indoor activities. As the GPS localization system is interfered by obstacles such as walls [1], indoor localization services cannot be realized. Indoor localization and navigation service LBS [2, 3] has become the focus of current market demand competition. The service quality of indoor LBS depends to a large extent on the localization accuracy of users [4]. In this context, Wi-Fi localization, Bluetooth localization, RFID localization, UWB (Ultra-Wideband) localization, infrared technology, ultrasonic and other technologies have entered the market one after another, contributing many effective location service solutions to the indoor localization needs of different industries.

Wi-Fi indoor localization technology [5, 6] is a relatively mature and widely used technology at present, and locates by collecting Wi-Fi signal values. The Wi-Fi localization technology is applied to small-scale indoor localization, and has low cost, but it is easily affected by many environmental factors. Bluetooth technology [7, 8] is not very different from Wi-Fi, and it is also greatly interfered by noise signals. However, its advantages are the small size, short distance, low power consumption, and easy integration in mobile devices such as mobile phones. The basic principle of RFID localization [9, 10] is to read the characteristic information of the target RFID tag through a set of fixed readers, and then determine the position of the tag through the nearest

* Corresponding author
E-mail address: classforyc@163.com

neighbour method, multilateral localization method, etc. However, RFID localization technology cannot achieve real-time tracking. UWB localization technology [11, 12] has strong multipath resolution and high accuracy, and the localization accuracy can reach sub-meter level. UWB pulse signals are analysed and positioned by multiple sensors using TDOA and AOA localization algorithms. However, its cost is relatively high and cannot be universally covered and used. Infrared technology [13] is used for indoor localization with high accuracy, but infrared can only spread through line-of-sight, with very poor penetration, so the cost is high and the localization effect is limited. Ultrasonic indoor localization technology [14, 15] mainly adopts reflective ranging method, with high localization accuracy and simple structure. However, it is greatly affected by multipath effects and non-line-of-sight propagation, and the frequency of the supersonic wave is affected by the Doppler effect and temperature.

However, most indoor localization research is based on a two-dimensional plane, that is, a single-story environment. For the localization of multi-storey buildings, only two-dimensional space technology cannot meet the demand. Since shopping malls, hospitals, airports, factories, etc. are all multi-story environments, floor localization is also crucial. At present, some methods have been proposed for floor recognition, such as Bayesian classifiers, artificial neural networks, and so on. However, some of these methods have high computational complexity and will affect the real-time performance of localization. Of course, some people directly use the barometer in the mobile phone for floor recognition. But as far as we know, quite a few mobile phones do not have a barometer. Therefore, we designed a two-story building recognition model based on Bluetooth fingerprint and LDA. In the offline phase, we collect the Bluetooth signal values of the two floors and train the floor recognition model through linear discriminant analysis. In the online phase, we first locate the floor through the trained model, and then locate the specific location using the improved KNN algorithm. Our main contributions are as follows:

We analysed the average RSS error between two floors of the same AP, and found that the classification model method to distinguish floors is feasible. We used the linear discriminant analysis method, and only selected 2 APs to complete the distinction between the two floors, and established a recognition model. The accuracy of floor recognition is 98% to 100%.

After localization the floor, we used the improved KNN algorithm to locate the specific location, and the localization accuracy was about 2.45m. Compared with the localization accuracy of 3.89m without floor localization, an increase of about 37%.

The model we designed has low cost, low computational complexity and high localization accuracy.

2. Method and Model

Our experiment is divided into two phases: offline phase and online phase. The offline phase is responsible for collecting data and training the model, and the online phase realizes the output of the target localization position.

2.1. Offline Phase

(1) Establish a local fingerprint database

Assume that N APs are placed in a multi-story building environment, and the signal values of these N APs can be received in all areas of all floors. In each floor, we set M sampling points to collect the RSS signal strength value of the AP. The locations of the sampling points on the upper and lower floors correspond to each other vertically, that is, the coordinates of the sampling points on each floor are the same. The collected RSS values are saved in the local database. The label structure of each sampling point in the database is $[F_i, G_j]$, where F_i represents the i-th floor and G_j represents the j-th sampling point. The fingerprint structure is $[RSS_1, RSS_2, \dots, RSS_n]$, where RSS_n represents the collected signal strength value of the n-th AP.

(2) LDA-based identification of two-story building

First, we deployed 4 APs in a two-story experimental environment, and measured the average error of the same AP between adjacent floors. The average RSS difference between adjacent floors of the same AP is between 9.25 dBm and 23.11 dBm. Therefore, based on the difference of RSS values in each floor, it is feasible to distinguish the number of floors through classification model. In this article, we use a classic method "Linear Discriminant Analysis" (LDA) as the floor classification method.

LDA is a supervised dimensionality reduction technology. The idea of LDA can be summarized in one sentence: after projection, the intra-class variance is the smallest, and the inter-class variance is the largest. It means to project the label data set at low latitudes. After projection, it is hoped that the projection points of the same category of data are as close as possible, and the distance between the category centre points of different categories of data is as large as possible. For example, for a two-dimensional feature data set, the two-dimensional data set is projected onto a straight line. For this reason, we derive and calculate the "generalized Rayleigh quotient" according to the idea of LDA. The result of the best classification is the projection result when the generalized Rayleigh quotient obtains the maximum, that is, the distance within the class is the smallest and the distance between the classes is the largest. Then when classifying the new sample data, the samples will be projected onto the same straight line. We determine the category based on the distance from the projection point to the centre point of the category. In the indoor multi-floor experiment, the greater the number of APs, the higher the computational complexity, so we choose the AP with the best classification performance as the attribute of floor classification.

For the application of this article, we use LDA for binary classification. In detail, suppose that in a two-story building, we deploy N APs and M sampling points on each floor. Our data set $D=\{(x_1,y_1),(x_2,y_2),\dots,(x_k,y_k)\}$, where any sample x_i is an N-dimensional vector, $y_i\in\{0,1\}$, $k=M\times N$. Since there are two types of data, we only need to project the data onto a straight line. Our projection line is a vector w, for any sample x_i and the category centre point $\mu_j(j=0,1)$, they are projected as $\omega^T x_i$ and $\omega^T\mu_j$ here. LDA needs to make the distance between the category centres of different categories of data as large as possible, so we maximize $\|\omega^T\mu_0-\omega^T\mu_1\|_2^2$. At the same time, we need to make the projection points of the same type of data as close as possible, that is, to minimize the covariance of the projection points of the same sample, which is to minimize $\omega^T\sum_0\omega+\omega^T\sum_1\omega$ ($\sum_i$ is the covariance matrix of the i-th sample). Therefore, we need to find a projected straight line so that it meets the above conditions. The formula is as follows:

$$arg\ \max_\omega J=\frac{\|\omega^T\mu_0-\omega^T\mu_1\|_2^2}{\omega^T\sum_0\omega+\omega^T\sum_1\omega} \quad (1)$$

where ω is the projection vector, μ is the category centre, and $\sum$ is the covariance matrix.

Transform (1) to get:

$$arg\ \max_\omega J=\frac{\omega^T(\mu_0-\mu_1)(\mu_0-\mu_1)^T\omega}{\omega^T(\Sigma_0+\Sigma_0)\omega} \quad (2)$$

We set the intra-class divergence matrix S_ω and the inter-class divergence matrix S_b:

$$S_\omega=\Sigma_0+\Sigma_1=\sum_{x\in X_0}(x-\mu_0)(x-\mu_0)^T+\sum_{x\in X_1}(x-\mu_1)(x-\mu_1)^T \quad (3)$$

$$S_b=(\mu_0-\mu_1)(\mu_0-\mu_1)^T \quad (4)$$

So we continue to transform (2) to get:

$$arg\ \max_\omega J=\frac{\omega^T S_b\omega}{\omega^T S_\omega\omega} \quad (5)$$

where J is called the generalized Rayleigh quotient. The larger the generalized Rayleigh quotient, the better the two categories can be distinguished. According to the properties of the generalized Rayleigh quotient, we get $S_\omega{}^{-1}S_b\omega=\alpha\omega$. Since the directions of $S_b\omega$ and $\mu_0-\mu_1$ are always parallel, let us set: $S_b\omega=\alpha(\mu_0-\mu_1)$. The two formulas are combined to obtain the calculation formula of ω:

$$\omega=S_\omega{}^{-1}(\mu_0-\mu_1) \quad (6)$$

LDA uses the data set D to find the projection w when J takes the maximum value to obtain the best floor two classification result.

Obviously, the number of APs is the number of sample attributes, which largely determines the amount of calculation for floor classification. Therefore, we decided to select only two APs as features. While reducing the amount of calculation, we found through experiments that the classification effect of 2 APs is also very good. There are currently N Aps. We can choose two different APs as a group, and there are $N\times(N-1)/2$ groups in total. According to the two APs of each group, calculate the best projection w and the generalized Rayleigh quotient J for this group. There are $N\times(N-1)/2$ groups ω and J. The larger the value of J, the better the classification effect of the two floors. Therefore, we select the group of APs with the largest J value as the characteristic attributes of the two floors, and the projection vector ω corresponding to this J is the best classification result of the two floors.

2.2. Online Phase

We measure the RSS value of the AP and input it into the floor classifier to obtain the floor location result. Then we perform specific location based on the improved KNN algorithm. We have n APs and m sampling points. The RSS sequence we measured is $[RSS_1,RSS_2,RSS_3,\dots,RSS_n]$. Find the Euclidean distance between this sequence and each coordinate point $[l_1,l_2,\dots,l_m]$. We set the three points with the smallest distance $(x_1,y_1),(x_2,y_2),(x_3,y_3)$, and the distances are l_1,l_2,l_3 respectively. From this, we calculate the weight:

$$\alpha_k=\frac{\frac{1}{l_k}}{\frac{1}{l_1}+\frac{1}{l_2}+\frac{1}{l_3}}\ k=1,2,3 \quad (7)$$

According to (7), we calculate the final localization result (x,y):

$$(x,y)=\alpha_1(x_1,y_1)+\alpha_2(x_2,y_2)+\alpha_3(x_3,y_3) \quad (8)$$

3. Experiment and Analysis

3.1. Experiment Environment

The experimental environment is the two-story building of our school. We selected 300 square meters for each floor, a total of 600 square meters as the experimental area. The adjacent sampling points of each layer are separated by 1m, that is, there are 4 sampling points on a $1m\times 1m$ grid.

There are 300 sampling points on each floor, and a total of 600 sampling points on two floors. Measure the RSS value of 1 minute at each sampling point, and then average the RSS value of this minute as the fingerprint of the sampling point. In actual measurement, not all areas can receive the signal value of all APs. Therefore, we set the RSS signal of the AP not received to -100dBm. At the same time, we found that the received RSS signal value is less than -100dBm, indicating that the AP signal is particularly weak. We set the value less than -100dBm to -100dBm.

3.2. Analysis of the Experimental Results of Floor Recognition

We first arranged 4 APs in two floors and measured the RSS values of these 4 APs at each sampling point. We train the model. Then we group the 4 APs in pairs, a total of 6 groups. For each group of AP, we calculate the generalized Rayleigh quotient J and its corresponding projection vector ω. Table 1 shows the values of the generalized Rayleigh quotient for 6 groups of AP.

According to Table 1, we can see that the combination of AP1 and AP4 has the largest generalized Rayleigh quotient. These two APs are the attributes of the optimal classification result.

Table 1. J between two Aps.

	AP2	AP3	AP4
AP1	32.94	35.34	49.89
AP2	X	42.23	24.39
AP3	X	X	28.68

Using the combination of AP1 and AP4, we randomly selected 90 test points on the two floors, a total of 180 test points. The test result is shown in Figure 1. There are three prediction errors, and the recognition accuracy is 98.33%.

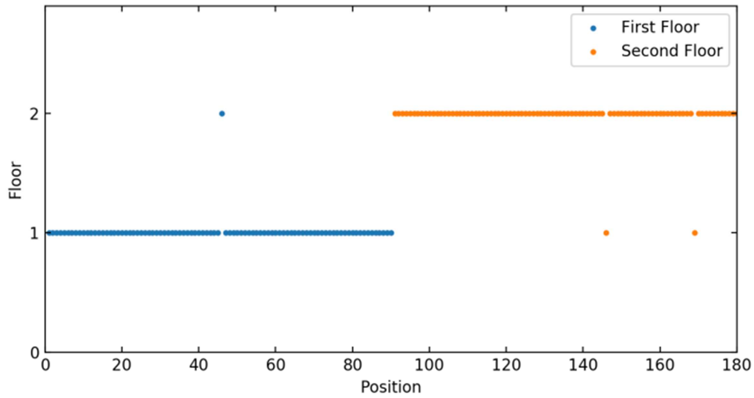

Figure 1. The result of test.

3.3. Localization Experiment Result Analysis

After obtaining the location of the floor, we continue to locate the specific location on that floor. We select the AP on the floor and use the improved KNN algorithm to obtain the specific location of the target. For each prediction point, we selected the nearest 3 sampling points from the map and gave each sampling point a weight. The localization result is shown in Figure 2. The average localization error after floor recognition is 2.45m, while the average localization error without floor recognition is 3.89m.

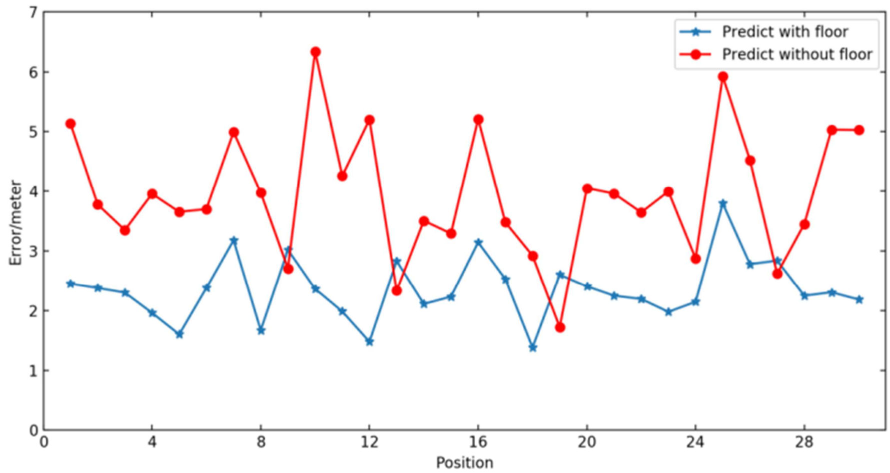

Figure 2. Comparison with or without floor localization.

4. Conclusion

In a multi-floor, if you want to know where you are, floor recognition is the first step in localization. Based on theoretical analysis and experimental testing, this paper proposes a two-story building recognition model based on LDA. We first group the arranged APs in pairs and calculate their generalized Rayleigh quotient J respectively, and then select the group of APs with the largest J value as the attributes of the two floors. Then, we collected the fingerprint data of the two floors, which proved that the model we built can reduce the calculation cost while ensuring the high accuracy of floor recognition. After determining the floor location, we use an improved KNN algorithm to obtain specific location results. We compare the localization accuracy with floor recognition and the localization accuracy without floor recognition. The experimental results show that the localization accuracy through floor recognition is higher and the calculation cost is lower.

Our research needs further expansion. When the floor area becomes larger, a small number of APs cannot meet the classification requirements. We need to increase the deployment of APs and the selection of attribute APs. When there are more floors, we need to build a new and more complex floor recognition model. In addition, we can also add motion recognition to it to trigger localization for real-time tracking when going up and down or taking an elevator.

References

[1] Jiung-yao Huang and Chung-Hsien Tsai, "Improve GPS positioning accuracy with context awareness," 2008 First IEEE International Conference on Ubi-Media Computing, Lanzhou, 2008, pp. 94-99.

[2] P. Bahl and V. N. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No. 00CH37064), Tel Aviv, Israel, 2000, pp. 775-784 vol. 2.

[3] Youssef M, Agrawala A. The Horus WLAN location determination system [J]. Wireless Networks, 2005, 14 (3): 357-374.

[4] Dimitrios Lymberopoulos, Jie Liu, Xue Yang, Romit Roy Choudhury, Vlado Handziski, and Souvik Sen. 2015. A realistic evaluation and comparison of indoor location technologies: Experiences and lessons learned. In Proc. IPSN. ACM, New York, NY, USA, 178–189.

[5] Hu Li, Yao-hui Wang, Qi-ming Sun and Jin-nan Liu, "Indoor localization for sparse wireless networks with heterogeneous information," 2015 International Conference on Information Networking (ICOIN), Cambodia, 2015, pp. 200-204.

[6] Ting-Ting Xiao, Xing-Yu Liao, Ke Hu and Min Yu, "Study of fingerprint location algorithm based on WiFi technology for indoor localization," 10th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM 2014), Beijing, 2014, pp. 604-608.

[7] X. Hou and T. Arslan, "Monte Carlo localization algorithm for indoor positioning using Bluetooth low energy devices," 2017 International Conference on Localization and GNSS (ICL-GNSS), Nottingham, 2017, pp. 1-6.

[8] S. G. Obreja and A. Vulpe, "Evaluation of an Indoor Localization Solution Based on Bluetooth Low Energy Beacons," 2020 13th International Conference on Communications (COMM), Bucharest, Romania, 2020, pp. 227-231.

[9] A. Buffi, P. Nepa and R. Cioni, "SARFID on drone: Drone-based UHF-RFID tag localization," 2017 IEEE International Conference on RFID Technology & Application (RFID-TA), Warsaw, 2017, pp. 40-44.

[10] M. Wang, B. Xue, W. Wang and J. Yang, "The design of multi-user indoor UWB localization system," 2017 2nd International Conference on Frontiers of Sensors Technologies (ICFST), Shenzhen, 2017, pp. 322-326.

[11] A. Motroni, A. Buffi and P. Nepa, "Localization of a mobile device equipped with an RFID reader," 2017 IEEE International Conference on RFID Technology & Application (RFID-TA), Warsaw, 2017, pp. 74-79.

[12] A. Poulose, O. S. Eyobu, M. Kim and D. S. Han, "Localization Error Analysis of Indoor Positioning System Based on UWB Measurements," 2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN), Zagreb, Croatia, 2019, pp. 84-88.

[13] Chunhan Lee et al., "Indoor positioning system based on incident angles of infrared emitters," 30th Annual Conference of IEEE Industrial Electronics Society, 2004. IECON 2004, Busan, South Korea, 2004, pp. 2218-2222 Vol. 3.

[14] S. Latif, R. Tariq, W. Haq and U. Hashmi, "INDOOR POSITIONING SYSTEM using ultrasonics," Proceedings of 2012 9th International Bhurban Conference on Applied Sciences & Technology (IBCAST), Islamabad, 2012, pp. 440-444.

[15] K. W. Kim and C. G. Lee, "Experimental feasibility study on using ultrasonic reflections for indoor human movement tracking," 2011 IEEE International Conference on Service-Oriented Computing and Applications (SOCA), Irvine, CA, 2011, pp. 1-3.