

# Fingerprint database construction method based on optimized attenuation factor model

Haiyan Lan<sup>1,a</sup>, Di Zhou<sup>1,b</sup>, Xuesong Ma<sup>2,c</sup>, and Eryue Liang<sup>1,d</sup>

<sup>1</sup> College of computer science and technology, Harbin Engineering University, Harbin, Heilongjiang, 150001, China.

<sup>2</sup> Heilongjiang Academy of Sciences, Harbin, Heilongjiang 150001, China

<sup>a</sup> lanhaiyan@hrbeu.edu.cn, <sup>b</sup> 409310423@qq.com, <sup>c</sup> 404361372@qq.com, <sup>d</sup> ley95@qq.com

**Keywords:** indoor location, fingerprint database, optimal attenuation factor model

**Abstract.** Aiming at the construction stage of off-line fingerprint database of fingerprint positioning technology based on WiFi, for the construction of traditional fingerprint database needs field acquisition of fingerprint signal widely, it is difficult to achieve in a large area, a method is proposed to construct the fingerprint database based on optimal attenuation factor model. Firstly, the traditional attenuation factor model was optimized, the free space model and the attenuation factor model were combined, and through the signal acquisition of several nodes, it realized the self-adaption of the path loss index in attenuation factor model. Secondly, the method of constructing the fingerprint database by using the optimized attenuation factor model was given. Finally, the optimization is confirmed by simulating with a variety of classical propagation model.

## 1. Introduction

In recent years, location-based fingerprint positioning technology has been paid more and more by virtue of its high efficiency, universality, low power consumption and high positioning accuracy. Location-based fingerprint positioning technology is divided into two stages: offline fingerprint library establishment phase and online positioning stage. The scientific nature of the fingerprint library construction directly affects the accuracy of the whole positioning, and the time spent on the construction of the fingerprint library occupies the vast majority of the positioning. Therefore, the construction of the fingerprint library in the offline stage needs to use the scientific and efficient way to ensure a short time to achieve high-precision positioning. The traditional fingerprint library is constructed using the RSSI method. The experimenter collects signals in a large area to be positioned to form a fingerprint information base. This method builds the most accurate fingerprint library, but consumes a lot of manpower and time. Because the density of the fingerprint node is proportional to the positioning accuracy, it is difficult to realize the traditional measurement method with hundreds of thousands of nodes in a relatively large environment. Many scholars have come up with a solution to this problem. Prashant Krishnamurthy presented a sub-region positioning method, according to the region of the grid to exclude invalid fingerprint nodes to reduce the fingerprint collection workload in the article Advanced Location-Based Technologies and Services[1]. Park put forward the organic construction scheme of fingerprint library, using user-based RSS acquisition to reduce the workload of collecting location fingerprints, and using clustering and Voronoi diagram to organize the development of this fingerprint library[2]. Chen L proposed an improved bimodal Gaussian distribution to reduce the workload of off-line data training[3]. Through the comparison of the existing methods, we can think of the method of constructing the fingerprint information base by using the signal propagation model, and calculate the fingerprint information by calculating the strength and distance of the AP node to construct the fingerprint information base.

## 2. Typical propagation model

### 2.1 Logarithmic path loss model.

Because of the influence of multipath propagation and non-line-of-sight propagation, it has been proved that the long-distance fading characteristic of the channel follows the lognormal distribution[4]. Logarithmic path loss model is

$$P_L(d) = P_L(d_0) + 10n \lg\left(\frac{d}{d_0}\right) + X_\sigma \quad (1)$$

$P_L(d)$  is the total loss strength from the beacon node to the target node.  $P_L(d_0)$  is the loss of near the beacon node, according to the size of the measurement area is different,  $d_0$  usually take 1m and 100m as a reference.  $n$  is the path loss index;  $d$  is the distance to be measured;  $X_\sigma$  is the normal random variable of the standard deviation  $\sigma$ . The path loss index  $n$  and the common experience value of  $X_\sigma$  are different with the change of environment.

### 2.2MK model.

MK model is characterized by taking into account from the beacon node to the target node signal transmission process by the wall and floor loss factors. The MK Model is

$$L_{pico} = L_0 + 10 * n * \log(d) + N_w * L_w + N_f * L_f \quad (2)$$

$L_{pico}$  is the total loss strength from the beacon node to the target node.  $L_0$  is the distance from the beacon near the propagation loss, usually set to 1 m,  $n$  for the path loss factor,  $d$  is the propagation distance.  $N_f$  and  $L_f$ , respectively, that the signal from the transmitter to the receiving side of the total number of walls through the walls and wall loss factor.

### 2.3Attenuation factor model.

The attenuation factor model is divided into single-layer building model and multi-storey building model. The single-layer attenuation factor model is[5]:

$$P_L(d) = P_L(d_0) + 10\lambda \log\left(\frac{d}{d_0}\right) + FAF \quad (3)$$

$\lambda$  represents the single-layer path loss coefficient, the path loss coefficient in different environments,  $FAF$  is the loss of building through the building. The loss of different buildings vary.

## 3. Attenuation Factor Model Optimization

Experiments show that the attenuation factor model is closer to the law of indoor wireless signal propagation, so this paper on the attenuation factor model in-depth exploration and optimization. The above mentioned single-layer attenuation factor model model is  $P_L(d) = P_L(d_0) + 10\lambda \log\left(\frac{d}{d_0}\right) + FAF$ , if the RSSI is the transmit power of the AP node,

$$RSSI - P_L(d) = RSSI - P_L(d_0) - 10\lambda \lg\left(\frac{d}{d_0}\right) - FAF \quad (4)$$

Let  $RSSI_{(d_0)}$  be a reference point to receive the signal strength,  $RSSI_{(d_0)} = RSSI - P_L(d_0)$ , the model will be deformed as

$$RSSI_{(d_n)} = RSSI_{(d_0)} - 10\lambda \lg\left(\frac{d}{d_0}\right) - FAF \quad (5)$$

### 3.1The Combination of Free Space Model and Attenuation Factor Model.

After deformation of formula (5), the model near the reference point to receive signal strength  $RSSI_{(d_0)}$  in a small environment usually take 1 meter at the signal strength. The traditional method requires the use of a signal collector to measure the received signal strength at a distance of 1 meter from the AP node, which increases the complexity of the positioning to a certain extent, and the distance control and the singular values have an indirect effect on the positioning accuracy.

Taking into account the environment set by the AP node, there is no obstacle in the range of 1 meter, the AP node to the reference point is close to the definition of free space, so the free space

propagation model is applied to the reference point signal strength acquisition. Let  $d$  be assigned a value of 1/1000 to obtain a reference point loss  $Ploss_{(d_0)} = 32.44 + 20\lg f_{MHz} + 20\lg\left(\frac{1}{1000}\right)$  (6)

Reference point received signal strength is

$$RSSI_{(d_0)} = AP\_Power + AP\_Gain + FP\_Gain - \left( 32.44 + 20\lg f_{MHz} + 20\lg\left(\frac{1}{1000}\right) \right) \quad (7)$$

This reduces the error caused by human factors, while increasing the universality to obtain the strength of reference point.

### 3.2 Path Loss Index and FAF Adaptive.

From the formula (5) we can see that the deformation of the model contains the path loss index  $\lambda$  and noise  $FAF$  two uncertainties. In the application of the traditional attenuation factor model, the experience values of different environments such as Table 2 and Table 3 are usually used, and because of the diversity of the environment, this can not be accurately positioned. When configuring the AP node, select a master node to obtain the signal strength from the master node when configuring other AP nodes. Since the relative coordinates of the nodes can be acquired when the AP nodes are configured, the node spacing can be calculated, as shown in Fig. 1.

In this case, formula (5) contains only the path loss index  $\lambda$  and the noise  $FAF$  two position variables. They can be determined by the two equations. Take AP0, AP1, AP2 as an example.

$$\begin{cases} RSSI_{(AP1)} = RSSI_{(d_0)} - 10\lambda\lg\left(\frac{d_{01}}{d_0}\right) - FAF \\ RSSI_{(AP2)} = RSSI_{(d_0)} - 10\lambda\lg\left(\frac{d_{02}}{d_0}\right) - FAF \end{cases} \quad (8) \quad \text{Come to a result} \quad \begin{cases} \lambda = \frac{RSSI_{(AP1)} - RSSI_{(AP2)}}{10 * \lg\left(\frac{d_{02}}{d_{01}}\right)} \\ FAF = RSSI_{(d_0)} - RSSI_{(AP2)} - 10n\lg\left(\frac{d_{02}}{d_0}\right) \end{cases} \quad (9)$$

However, the number of AP nodes that need to be configured for each location will vary. When the AP node is more than three, in order to make full use of AP node information, greatly improve the path loss index  $\lambda$  and noise  $FAF$  accuracy. The corresponding path loss exponent  $\lambda$  and noise  $FAF$  are calculated by combining any two AP nodes except the main AP node.

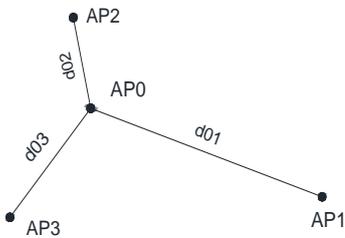


Fig. 1. AP node diagram

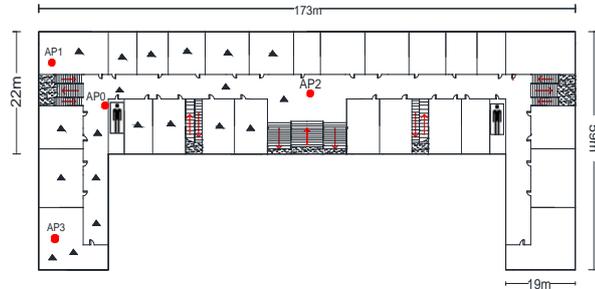


Fig.2 21a third floor simulation figure

### 3.3 Multi model simulation

Select the third floor of the 21a teaching building of Harbin Engineering University as the experimental site, the simulation map is shown in Fig. 2. The solid origin in the figure is AP node, and the triangle solid point is the field detection location. Because of the large size of the experimental site, considering the transmission power of WiFi and the resources, half of the selected map as the experimental site. There are four AP nodes on the site, AP0 (CMCC-EDU) is the main AP node, The transmit power is 20dbm and the gain is 7dbi. The remaining three AP nodes transmit power of 18dbm, gain of 6dbi. The lower left corner of Fig.2 is set as the origin of coordinates to determine AP node coordinates, AP0 (18,35), AP1 (2,56), AP2 (84,46), AP3 (2,2). Thus we can get the distance between three AP nodes and the main node are  $d_{01}$ ,  $d_{02}$  and  $d_{03}$ . When the AP node is configured, the signal intensity of three AP nodes receiving the main AP node are AP1 (-60dbm, 26.4m), AP2 (-74dbm, 66.9m) and AP3 (-72dbm, 36.7m).  $\bar{\lambda}$  and  $\overline{FAF}$  are calculated according to formula (10) are 2.57 and 21.93. At the same time, the twenty experimental sites were selected to test the model, and

the results were compared with the empirical attenuation factor model, the log path loss model, the MK model and the free space propagation model. The receiving gain of notebook computer for signal acquisition is 2dbi. The signal intensity distribution of the five models is shown in Fig. 3.

It can be seen from Fig.3 that the free space model does not consider the effects of multipath propagation and NLOS propagation, and the result is much higher than the actual value. The result of MK model is obviously higher, because the MK model mainly focuses on the relationship between the loss of the wall and floor, the empirical value of the path loss index is 2, which is smaller than other models. However, the experimental site is not multilayer, and the floor loss is ignored, so the calculated value will be high, which proves that the MK model is not suitable for the construction of the indoor fingerprint database. After the optimization, the attenuation factor model has obvious improvement compared with the empirical attenuation factor model, However, in the distance from the AP node 20m, the strength of the test is not low, The log path loss model is more accurate in this range, but with the increase of distance and the increase of obstacles such as walls, the optimized attenuation factor model has a high fitting degree with the actual measured values. The CDF error curve of the simulation data is shown in Fig. 4.

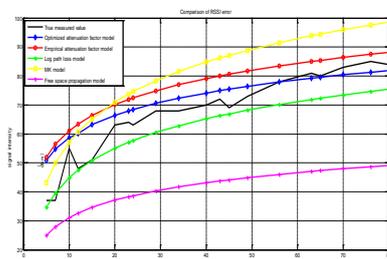


Fig. 3 Comparison of distance and signal intensity

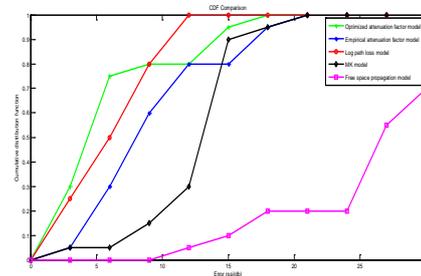


Fig. 4 error curve of CDF

#### 4. Summary

This paper presents the optimization attenuation factor model includes two parts: With the combination of free space propagation model and attenuation factor model, the attenuation of the reference point in the model by using the free space loss model that reduces the measurement time and the error of human operation and the universality of the algorithm is increased. According to the measurement, different path loss index and FAF are used to adapt to the environment. Compared with the traditional attenuation factor model with the empirical value of the path loss exponent, it has higher accuracy in different positioning environments. In this paper, by contrast with the four typical models, confirmed the optimized attenuation factor model in different size space and different environment has very high accuracy, compared with the traditional empirical attenuation factor model has greatly improved. In the large site and AP nodes near the barrier free space, the value calculated by optimal attenuation factor model is on the low site, and worse than log path loss model. The log path loss model can be taken into account in the near range, and the optimal attenuation factor model is used at the far end, the combination can achieve better results.

#### 5. Acknowledgements

This work was financially supported by the Heilongjiang Youth Natural Fund (QC2014C067), Heilongjiang Postdoctoral Research Foundation (LBH-Q14056), Harbin Science and Technology Innovation Fund 2015RQQXJ072, 2016RQQYJ091) and Central University Fund(HEUCS201702).

#### References

- [1] Karimi H. Advanced location-based technologies and services[J]. Taylor & Francis, 2013.
- [2] Park J G, Charrow B, Curtis D, et al. Growing an Organic Indoor Location System[C]// International Conference on Mobile Systems, Applications, and Services. 2010:271-284.

- [3] Chen L, Li B, Zhao K, et al. An Improved Algorithm to Generate a Wi-Fi Fingerprint Database for Indoor Positioning[J]. *Sensors*, 2013, 13(8):11085.
- [4] Jung S, Lee C O, Han D. Wi-Fi fingerprint-based approaches following log-distance path loss model for indoor positioning[C]// *IEEE Mtt-S International Microwave Workshop Series on Intelligent Radio for Future Personal Terminals*. IEEE, 2011:1-2.
- [5] Shih Y, Hsu Y, Chen C, et al. Adaptive attenuation factor model for localization in wireless sensor networks[J]. *International Journal of Pervasive Computing and Communications*, 2008, 4(3):257-267.