

# Study of WLAN Fingerprinting Indoor Positioning Technology based on Smart Phone

Ye Yuan <sup>a</sup>, Daihong Chao, Lailiang Song

School of Instrumentation Science and Opto-electronics Engineering, BeiHang University, Beijing  
100191, China

<sup>a</sup>buaayuanye@163.com

**Keywords:** indoor positioning; wireless LAN fingerprint; pattern matching; smart phone.

**Abstract.** Wireless LAN fingerprinting positioning technology is applied in indoor space equipped with some Wireless LAN access points. WLAN fingerprinting positioning technology utilize specified devices to receive WLAN signal and analyze signal strength in order to establish complete fingerprint database which experimental data can be loaded into to accomplish final positioning target under the instruction of specified pattern matching algorithm. Smart phone was developed into such a device to receive WLAN signal. After establishing fingerprint database, two pattern matching algorithms are used to achieve positioning target and verify the influence on positioning accuracy by WLAN signal direction and WLAN access point number.

## Introduction

Currently indoor positioning technology<sup>[1][2][3]</sup> based on WLAN signal is a major research focus within the scope of navigation and positioning technology which has the potential for enormous practical value. Current indoor positioning algorithms are mainly inherited from the traditional outdoor positioning technology, including TOA (Time of Arrival), AOA (Angle of Arrival), TDOA (Time Difference of Arrival) etc<sup>[4][5]</sup>. Nowadays with the widespread popularity of IEEE802.11 protocol, positioning algorithms based on WLAN signal strength gradually play an important role in indoor positioning technology<sup>[6]</sup>. Compared with the input parameter of traditional algorithms, such as the arrival time and arrival angle, the signal strength is more easily available and the accuracy can also be better assured. This advantage greatly enhances the operability of RSS (Received Signal Strength) positioning algorithm and laid solid foundation for RSS algorithm wide utilization in ordinary portable terminal devices, such as smart phones, tablet PCs and other wearable device.

To implement RSS algorithm, experimenter should go through the following flow. Firstly experimenter divide space installed with several wireless networks AP (Access Point) into a scene positioning grids, each grid center called signal CP (Calibration Point). Then the signal strength is measured by the value of all APs consisting of WLAN signal strength vector at each CP by the portable device in order to establish a complete fingerprint database (Fingerprinting). Finally the signal strength is sampled at each TP from all APs and loaded into the above Fingerprinting database in order to obtain the final positioning estimated value utilizing specified pattern matching algorithm.

This article is presented as the following order: pattern matching algorithm is introduced as chapter 1 and then experimental condition is introduced in chapter 2. In chapter 3, not only basic positioning result with the above pattern matching algorithm but also additional detailed study about influence on positioning precision from some chosen factors such as WLAN signal directional issue, Access Point Numbers and category are presented. Chapter 4 contains summary about all the contents in this article.

## Introduction of Pattern Matching Algorithm

**K-Nearest-Neighbor Algorithm.** As the fundamental pattern classification algorithms, the core idea of K-Nearest-Neighbor algorithm is derived from Euclidean theorem, which calculates the distance between sample and all known patterns and estimate the sample as the known pattern with the

shortest distance. Further, in order to improve the classification accuracy, the estimated value of sample is usually closest to the mean of K known patterns.

In the calibration phase, each node receives the AP CP signal strength vector formula:

$$\bar{z}_\ell = [z_{\ell,1}, \dots, z_{\ell, N_{AP}}] \quad (1)$$

Among them,  $\ell = 1 \dots N_{CP}$

$\bar{z}_{i,j}$  Characterizes at position  $i$ , signal strength value from access point  $j$ ;

$N_{AP}$  Characterizes the number of access points in the specified space;

$N_{CP}$  Characterizes the number of calibration points in the specified space

In the estimation stage, each CP signal strength vector from all AP values is shown by the following formula:

$$\bar{y} = [y_1, \dots, y_{N_{AP}}] \quad (2)$$

Within the signal strength space, Euclidean distance between CP and AP is shown as following:

$$d(y, z_\ell) = \sqrt{\sum_{j=1}^{N_{AP}} (y_j - z_{\ell,j})^2} \quad (3)$$

Location can be estimated as the nearest neighbor in the signal space. Assume that  $k = 1$ ,

$$\hat{\ell} = \arg \min_\ell (d(y, z_\ell)) \quad (4)$$

In order to improve the classification accuracy and keep it universal, generally the average of K known patterns with the closest distance are taken as the estimation of the sample

In this paper, we assume  $K = 3$ .

**Kernel density estimation algorithm.** In probability theory, kernel density estimation is used to estimate the unknown density function as one of the non-parametric test methods. Since the kernel density estimation method does not need any prior knowledge and assumption of the data distribution and just obtains distribution characters from data sample itself, it attracts high degree of attention both in statistical theory and applications area.

According to Bayesian theory, the location estimation of the probability density function is shown as following formula:

$$p(\ell|y) = \frac{p(y|\ell)p(\ell)}{p(y)} = \frac{p(y|\ell)p(\ell)}{\sum_{\ell' \in \xi} p(y|\ell')p(\ell')} \quad (5)$$

Among them,

$p(\ell)$  Characterizes the prior probability density of the estimated position for position  $\ell$

Collection  $\xi$  contains all the possible positions;

$p(y)$  Characterizes the probability of the signal strength vector  $y = [y_1, \dots, y_{N_{AP}}]$  present in all possible positions.

According to Bayesian law, the probability function to the maximum likelihood estimate represents estimated value of TP shown as following:

$$\hat{\ell} = \arg \max_\ell p(\ell|y) \quad (6)$$

Considering the prior probability density  $p(\ell)$  are equal for all CP position, and  $p(y)$  has nothing to do with CP position, formula (5) and (6) can be simplified as following:

$$\hat{\ell} = \arg \max_\ell p(\ell|y) = \arg \max_\ell p(y|\ell) \quad (7)$$

Considering the independence of signal strength, namely maximum likelihood estimation (7) can be reduced to the following formula:

$$\hat{\ell} = \arg \max_\ell \left( \prod_{i=1}^{N_{AP}} p(y_i|\ell) \right) \quad (8)$$

$p(y_i|\ell)$  Is the kernel function needed to be estimated? Common kernel functions include the window function, the polynomial function, the Gaussian function etc. Taking into account of the intensity distribution of radio signal in space accordance with normal distribution, we define Gaussian function as kernel function as following formula shown:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (9)$$

That is

$$p(y_i|\ell) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - z_{\ell,i})^2}{2\sigma^2}\right) \quad (10)$$

## Introduction of Experimental Condition

Portable device used in this test to sample WLAN signal strength is Meizu smart phone. By calling built-in API framework agreement with specified Andriod program, radio signal source name SSID and respective signal strength can be recorded and saved in real time, preparing for subsequent algorithm analysis.

The real experiment is performed in Beijing University of Aeronautics and Astronautics New Main building, meshing within 40m \* 10m space and selecting 48 CP and 20 TP. As shown in Figure 1.

In 48 CPs experimenter samples WLAN signal strength in four directions however in 20 TPs experimenter randomly choose one direction to receive WLAN signal.

The total number of active WLAN access point in this experiment is 6, i.e.  $N_{AP_{max}} = 6$

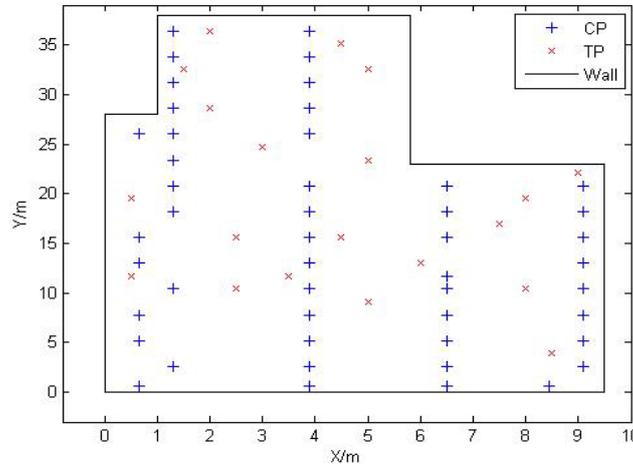


Fig 1 Distribution maps of experimental space

## Experimental Result Analysis

**Positioning algorithm implementation and comparison.** For 3-nearest neighbor algorithm, each CP signal strength vector from all the APs is shown as following:

$$\vec{z}_\ell = [z_{\ell,1}, \dots, z_{\ell, N_{AP}}] \quad (11)$$

Indeed, each point can be obtained 4 CP RSS vectors which are named with  $\vec{z}_{\ell E}, \vec{z}_{\ell W}, \vec{z}_{\ell S}, \vec{z}_{\ell N}$

Excluding the RSS direction, the final results from the CP points RSS is defined as the average of the four directions characterization:

$$\vec{z}_{\text{aver}} = \frac{1}{4} \sum (\vec{z}_{\ell E} + \vec{z}_{\ell W} + \vec{z}_{\ell S} + \vec{z}_{\ell N}) \quad (12)$$

Taking  $\bar{z}_{\ell_{aver}}$  into equation (3) for calculation and using equation (4) for determination.

For Gaussian kernel algorithm, we selected kernel function shown as equation (10).

Considering actual experiment condition, WLAN signal strength output are integers, which implies the RSS step equal to 1. So we can assume that standard deviation  $\sigma = 1$  and therefore the maximum likelihood estimation in the form of Gaussian kernel function depends on  $z_{\ell,i}$ .

Similar to the K-nearest neighbor algorithm, if not considering RSS directivity,  $z_{\ell,i}$  is given by the following formula:

$$z_{\ell,i} = \frac{1}{4} \sum (z_{\ell E,i} + z_{\ell W,i} + z_{\ell S,i} + z_{\ell N,i}) \quad (13)$$

Taking  $z_{\ell,i}$  into equation (10) and the kernel function can be calculated then using equation (8) for the estimated position.

The calculation is shown as Fig 2. We can conclude from Fig 2 that both 3 Nearest Neighbor and Gauss Kernel Function can be used in indoor positioning and with similar positioning variance.

**Influence of directionality of WLAN signal on positioning accuracy.** When WLAN signal spreading in space, it will inevitably be obstructed by barriers and its signal strength will gradually descend. So through in the same position, experimenter sample WLAN signal in different direction then RSS will have a significant difference. As mentioned earlier, in this experiment we collect signal strength in four directions on 48 CPs, and collect signal strength in a randomly chosen direction on 20 TPs. Therefore the issue of directionality of WLAN signals is introduced in the algorithmic level of which we try to verify its effect in this chapter.

For 3-nearest neighbor algorithm, if considering directionality of WLAN signal, RSS on 48 CPs from four directions are treated as independent variables, so the number of vectors RSS vector  $\bar{z}_{\ell_{dir}}$  of CPs is equivalent to 192. Taking all of the  $\bar{z}_{\ell_{dir}}$  into the calculation equation (3) and the final estimated value is calculated by equation (4).

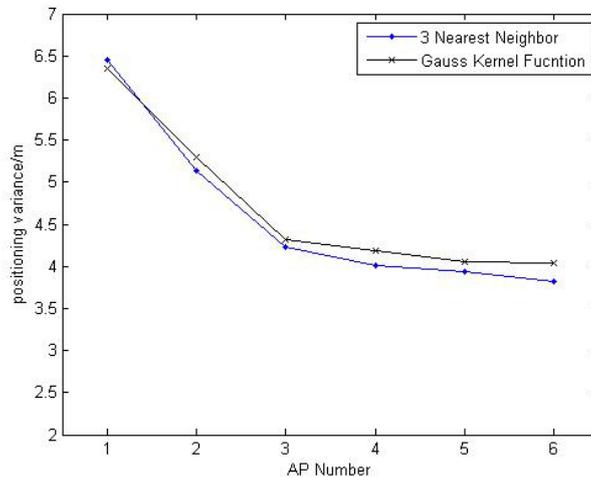


Fig 2 Positioning result based on two pattern matching algorithms

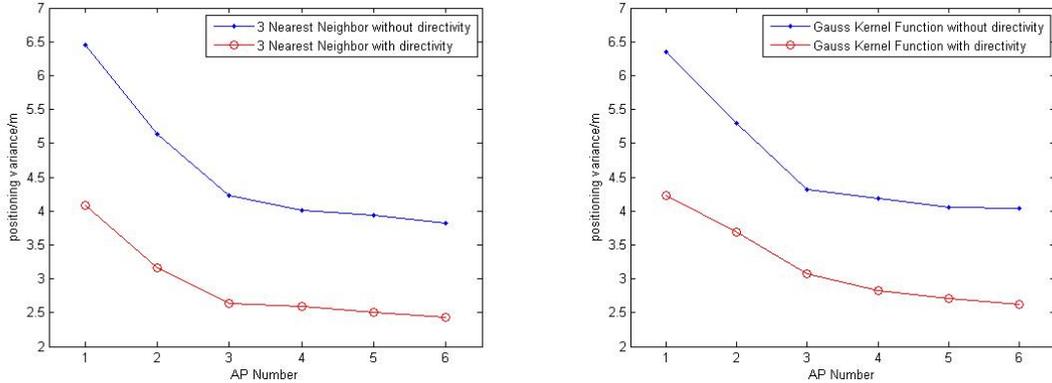
If we consider the directivity RSS for Gaussian kernel function, similar to above mentioned methods, RSS from four directions are treated as independent variable so single CP is equivalent to kernel function superposition of four directions, further decomposition of the equation (10):

$$p(y_i | \ell) = \frac{1}{4\sqrt{2\pi}\sigma} \sum \exp\left(-\frac{(y_i - z_{\ell E,i})^2}{2\sigma^2}\right) + \exp\left(-\frac{(y_i - z_{\ell W,i})^2}{2\sigma^2}\right) + \exp\left(-\frac{(y_i - z_{\ell S,i})^2}{2\sigma^2}\right) + \exp\left(-\frac{(y_i - z_{\ell N,i})^2}{2\sigma^2}\right) \quad (14)$$

Taken all the RSS value on four directions into equation (14), the kernel function can be calculated and then using the formula (8) to obtain the final estimated position.

The calculation result is shown as Fig 3.

According to Fig 3, if considering WLAN signal directivity, the precision variance of both algorithm will be significant descend. Therefore for all the calculation part involved in the latter part, algorithm structure will be utilized as chapter 3.2 shown.



(a) Calculation result of 3 nearest neighbor (b) Calculation result of Gauss kernel estimation

Fig 3 The Impact on positioning precision from WLAN signal direction

**Influence of number of access points WLAN signals on positioning accuracy.** As mentioned above, CP intensity RSS vector is described as  $\vec{z}_\ell = [z_{\ell,1}, \dots, z_{\ell, N_{AP}}]$ , TP RSS intensity vector is described as  $\vec{y} = [y_1, \dots, y_{N_{AP}}]$ .  $N_{AP}$  Represents the number of WLAN access points. More access points imply more information the intensity vector contains. In this paper  $N_{AP_{max}} = 6$ .

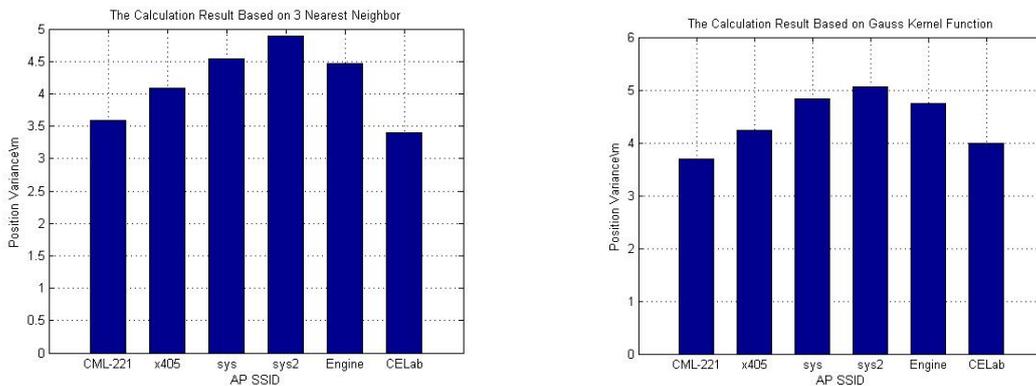
The relationship between the number of AP and the estimated position accuracy is shown as Figure 2 and Figure 3. Based on the analysis of above figure, we can get following conclusions:

Overall, number of AP implies the lower positioning error and the higher precision.

If the number of AP in space equals to 3 or less, the increase in the number of signal sources can significantly improve the positioning accuracy; however, when the number of AP is greater than 4, the increase of the number can only improve the positioning accuracy with very little small amplitude.

**Influence of type of WLAN access point signal on positioning accuracy.** Considering that six different locations in the spatial distribution of effective AP and the signal attenuation in the process of moving in different direction, we try to go deep into the micro-level to study whether the effect of each signal access point respectively generated on the positioning accuracy has obvious differences. In order to verify the difference, we will increase the AP number one by one with specified rules and then use such AP to estimate position in space with the above two algorithms respectively.

**Single AP for Positioning.** Assuming such a condition that there is only one AP in space, which is equivalent to use one of the above six selected AP signal source for independent solver. The results can be characterized by a respective signal quality of these AP which will impact on the positioning accuracy.



(a) Calculation result of 3 nearest neighbor (b) Calculation result of Gauss kernel estimation

Fig 4 Positioning Result based on signal AP

As shown in Fig 4, we can see the impact of different access points for positioning accuracy has obvious differences. Depending on the signal quality of the access point for position estimation, we can divided these six AP into two groups, which implies the best three signal quality of AP for one group, and the remaining three AP of the other group shown as following:

Good Quality Group: CML-221, x405, CELab;

Poor Quality Group: sys, sys2, Engine;

Dual APs for Positioning. In order to verify the influence of the different types of signal sources on positioning accuracy when the AP number equals to 2, we performed comparative experiments shown as following:

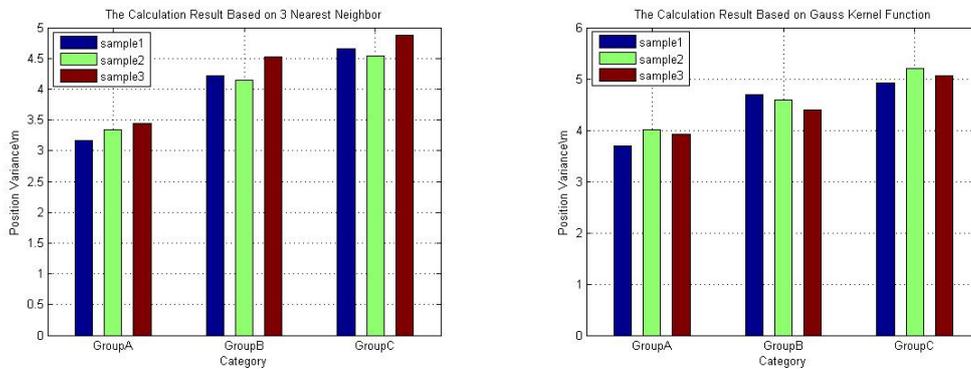
Group A: select two AP randomly from good quality group

Group B: select one AP randomly from good quality group and one AP randomly from poor one

Group C: select two AP randomly from poor quality group

In order to ensure the experiment result reliability, we perform above test three times named as sample1, sample2 and sample3 respectively.

Experimental result is shown as Fig 5. As shown in Fig 6, group A is better than group B, finally group C in positioning precision. The calculation result proves the assumption about group clarification.



(a) Calculation result of 3 nearest neighbor (b) Calculation result of Gauss kernel estimation  
Fig 5 Positioning Result based on two APs

Triple APs for Positioning. When we assume the number of AP equals to 3, similar to the above experimental ideas, we can make the following comparison test:

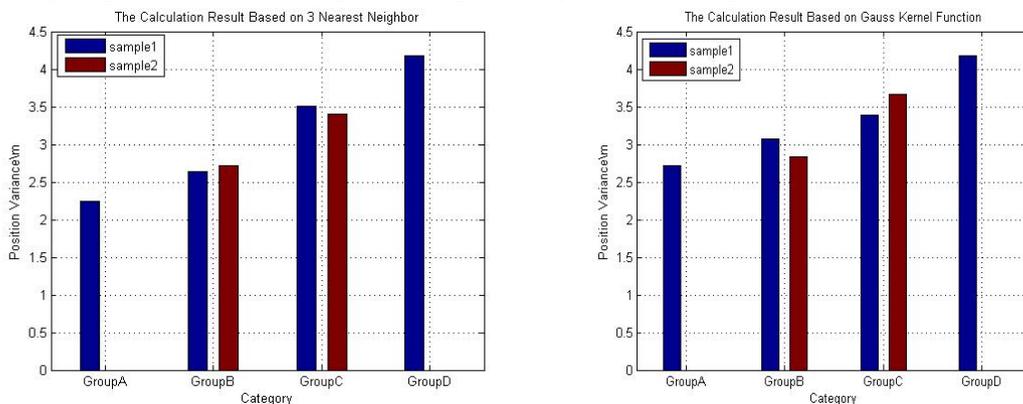
Group A: select all three AP in good quality group;

Group B: select two AP randomly from good quality group and one AP randomly from poor one;

Group C: select one AP randomly from good quality group and two AP randomly from poor one;

Group D: select all three AP in poor quality group;

Also, In order to ensure the experiment result reliability, we perform above test two times for group B and group C named as sample1, sample2 respectively. Experimental result is shown as Fig 6.



(a) Calculation result of 3 nearest neighbor (b) Calculation result of Gauss kernel estimation  
Fig 6 Positioning Result based on three APs

As shown in Fig 6, for positioning precision group A is better than group B, then group C and finally group D. The calculation result proves the assumption about group clarification.

According to the above experiments, we can recognize that different APs will have different impact on positioning accuracy because of the different locations in the actual space, different degrees of obstruction by the propagation process, different degree of attenuation and thus different signal quality which is marked for positioning.

## Summary

This article is based on specified physical space and figure prints database built by smart phone Android program, accomplishing and comparing two classical pattern matching algorithm in space positioning area with relative high position accuracy, further verifying the impact on positioning accuracy of multiple relative factors such as WLAN signal directionality, WLAN source number and category and thus laying a solid foundation for latter study about WLAN signal strength positioning algorithm.

## Reference

- [1] Seong Yun Cho, Chan Gook Park, Gyu In Jee, Measurement System of Walking Distance Using Low cost Accelerometers, The 4th Asian Control Conference, Singapore, Professional Activities Center National University of Singapore, 2002: 1799~1803.
- [2] C. Goodall, Z. Syed, N. El-Sheimy, A Truly Portable, Low-Cost, And Accurate Mobile Navigator For Urban And Indoor Usage, 23rd International Technical Meeting of the Satellite Division of The Institute of Navigation, Portland, OR, September 21-24, 2010:2969~2976.
- [3] S. Saeedi, Dr. N. El-Sheimy, X. Zhao, Dr. Z. Sayed, Context Aware Mobile Personal Navigation Using Multi-level Sensor Fusion, 24th International Technical Meeting of the Satellite Division of The Institute of Navigation, Portland, OR, September 19-23, 2011:1394~1403.
- [4] Z. Syed, J. Georgy, C. Goodall, M. M. Atia, N. El-Sheimy, Trusted Portable Navigator for Environment and User Independent Positioning, 24th International Technical Meeting of the Satellite Division of The Institute of Navigation, Portland, OR, September 19-23, 1447~1452.
- [5] Dietrich Brunn, Uwe D. Hanebeck, Hui Wang Andrei Szabo, Joachim Bamberger, Performance Comparison of Nonlinear Filters for Indoor WLAN Positioning, Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute of Computer Science and Engineering Universitat Karlsruhe (TH), Germany, Learning Systems Information and Communications, Corporate Technology Siemens AG, Munich, Germany.
- [6] S J Julier, J K Uhlmann. A general method for approximating nonlinear transformations of probability distributions. Technical Report, Robotics Research Group, Department of Engineering Science, University of Oxford, 1996.