# A New RSS Distance Calculation Algorithm Based on Tree-ring Distance in APs Rich Indoor Localization Environments

Baofeng Wang[1], Liming Chen[2], Zumin Wang[3], Mengmeng Xu[1], Shuai Tao[3]

[1] School of Network Engineering, Zhoukou Normal University, China
[2] School of Computing, University of Ulster, UK
[3] College of Information Engineering, Dalian University, China
baofengw@yeah.net, l.chen@ulster.ac.uk, wangzumin@163.com, 568062616@qq.com, taoshuai@live.cn

## Abstract

The fingerprint-based localization technique is one of the most popular indoor localization technologies. There are quite a few localization algorithms that use the RSS distance of position pairs to characterize their physical distance. In this paper, we introduce two coefficients to measure the relationship between RSS distance and physical distance. Based on the definition of tree-ring distance, we found that the characterization capability of RSS distance to physical distance is closely related to APs' tree-ring distance. To exploit this, through an in-depth analysis of the relationship between tree-ring distance and physical distance, we pointed out that the APs sets composed of APs at the edge positions of the positioning area makes the RSS distance better to characterize the physical distance. Further, we proposed a novel RSS distance calculation algorithm based on the comparison of tree-ring distances. In the algorithm, for each pairwise position, the abnormal APs are eliminated by the Mean+3S method, and the APs with larger tree-ring distance are selected to participate in the calculation of RSS distance, namely, for different pairwise positions, different APs subsets of all APs are selected to participate in RSS distance calculation. We evaluate the algorithm in a simulation study and initial results show that an APs set with 3 APs is sufficient to guarantee very strong correlation (the correlation coefficient>0.8) and very high consistency (the consistency coefficient>0.8) between RSS distance and physical distance, which demonstrates the effectiveness and the practicability of the algorithm.

**Keywords:** RSS, Indoor localization, Fingerprinting localization

## 1 Introduction

In the past decade, with the extensive development of wireless technology, the urgent need for location-based services (LBSs) has also increasingly accelerated, such as personal navigation, LBS delivery, medical services, telehealth, etc. Indoor localization techniques have received considerable attention both in the academic and industrial areas [1-3]. So far, various indoor localization techniques have been developed based on different radio communication technologies, such as wireless local area networks (WLANs) [4], acoustic signals [5], UWB [6], RFID [7], etc. Also, there are localization techniques that fuse two or more technologies, also called signal of opportunity based localization [7-8]. However, due to the complexity of the indoor environments and the increasingly high requirements for localization accuracy, indoor localization techniques are facing more and more challenges [9].

Of all the localization technologies, WLAN-based indoor localization is one of the most popular topics. According to the localization principle, it can be classified into localization techniques based on the propagation model [10], and localization techniques based on fingerprint [11]. For the former, the exact positions of access points (APs) are needed to know in advance, which is unrealistic in most indoor environments. While the latter consists of two stages, the offline stage samples RSS values at reference points (RPs) to build a fingerprint database, and the online stage locates a user by matching his RSS measurements with fingerprint through the localization algorithm [12]. It does not need to know the positions of APs, and the localization accuracy is higher [13]. In addition, no need to install additional infrastructure, low cost and easy deployment are also advantages of fingerprint-based localization technique.

Among the entire fingerprint matching algorithms, many ones that need to consider the RSS differences of position pairs, such as KNN, MDS, etc. These RSS differences are Euclidean distances between RSS vectors from all APs, so the status of each AP will affect the ability of RSS differences to characterize the physical distances. However, with the rapid development of wireless technology and the prevalence of the internet of things, especially in the context of smart city construction, it becomes easier than ever

*Corresponding Author: Baofeng Wang; E-mail: baofengw@yeah.net
DOI: 10.3966/160792642020092105017

before for smart devices to receive high-quality signals from increasing APs in public indoor environments. Hence, utilizing so abundant APs to build a fingerprint database and locate all positions of the indoor area will lead to the heavy workload at the offline stage and inevitable localization errors at the online stage, which is not reality.

In this paper, we term the ability of the RSS difference from an AP (or APs) to characterize the physical distance of a specific pairwise position as characterization capability. Our intuition is choosing appropriate APs sets can not only reduce the calculation intensity and complexity but also improve the characterization capability of RSS distance, which inspires us to carry out a quantitative analysis for the relationship between RSS distance and physical distance. We uncover that APs at different positions have different characterization capabilities to the physical distance of pairwise position subject to signal propagation laws. Abnormal APs and environmental changes also lead to inaccurate APs' characterization capability.

With the above factors, we present an algorithm to calculate the RSS distance of a specific position pair in the positioning area. Through the algorithm, we can calculate the RSS distance with stronger characterization capability for any pairwise position, to provide a key foundation for the localization algorithm based on RSS distance. The focus of this paper is to provide an optimized middleware between the fingerprint database and fingerprint matching algorithm for a kind of RSS distances based localization technique. The framework of the fingerprint-based indoor localization scheme and the role of our work (in the red box) is as shown in Figure 1. In the rich APs indoor application scenarios, all APs have the following features:

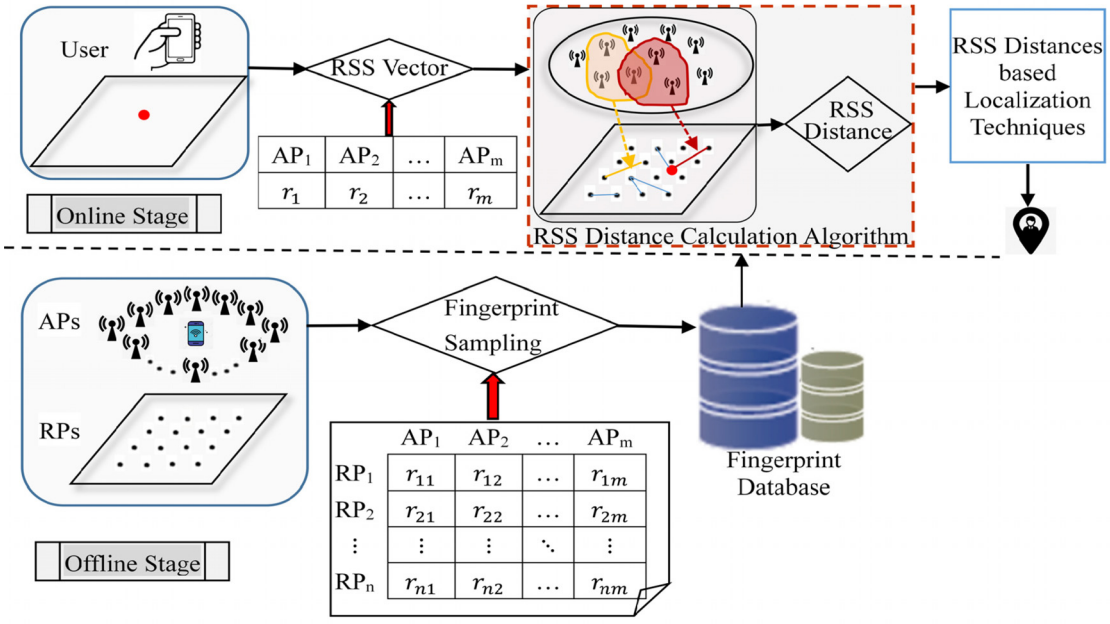



**Figure 1.** The framework of fingerprint-based indoor localization scheme and our main work

- They are deployed primarily for communication.
- They are not centralized management and debugging, or even installed by different individuals.
- They are intensive and well-distributed, for high-quality Wi-Fi signals are almost everywhere indoors.
- Their positions are unknown, and they do not communicate with each other.
- There are one or some APs that do not affect one position pair but are still effective for other position pairs, or there are some abnormal APs.

The main contributions of the paper are summarized as follows:

(1) To measure the relationship between RSS distance and physical distance, a correlation coefficient and a consistency coefficient are introduced.

(2) We first propose the concept of tree-ring distance of pairwise position orienting an AP, which can express the contribution of the APs at different positions to characterize the RSS distance.

(3) For each pairwise position, a new algorithm for RSS distance calculation is designed based on RSS tree-ring distance. In the algorithm, the abnormal APs are eliminated through the Mean+3S method, and the APs set participating in the calculation of RSS distance obtained by comparing the tree-ring distance.

(4) We point out that compared with an APs set with APs from the center region, the edge region makes the RSS distance better characterize the physical distance, which provides the guidance for APs deployment.

(5) By simulation study, we found that only 3 APs are enough to achieve a very strong correlation and a very high consistency between RSS distance and physical distance.

The rest of the manuscript is organized as follows. Section 2 reviews the related work on the two stages of the fingerprint-based localization. Section 3 presents

our preliminary knowledge and introduces two coefficients for measuring the relationship between RSS distance and physical distance, and in section 4, observation and analysis of the relationship between RSS distance and physical distance is depicted in detail. A new RSS distance calculation algorithm is detailed in section 5, followed by the experiments and simulation in Section 6, and the conclusions of the paper in Section 7.

## 2 Related Work

The localization accuracy of fingerprint-based localization has long been the primary challenge, which depends on the built fingerprint databases and the fingerprint matching algorithm. Over the last decade, a lot of schemes have been proposed to reduce the intensity of fingerprint collection and to improve the accuracy of fingerprint matching algorithms.

Based on the participation of ordinary users in data collection, a crowdsourcing strategy has been developed to reduce the work intensity of building a fingerprint database [14-15]. However, since the sample data contains no location label, the additional data of inertial sensor in the intelligent terminal is often used to estimate the unknown location label, which simultaneously causes issues in terms of accuracy, applicability, equipment heterogeneity, the additional power consumption of the sensor [14]. To reduce work intensity and improve accuracy, in [15], a LAAFU system has been proposed by employing implicit crowdsourced signals for fingerprint updates. In response to the heterogeneity of devices and the requirement of users' intervention in traditional crowdsourcing, researchers in [16] proposed a new strategy that incorporated a new preprocessing method for RSS samples, an implicit crowdsourcing sampling technique, and a semi-supervised learning algorithm. The strategy reduced the demand for large quantities of labeled samples and achieved good localization accuracy. Also, to remedy periodical calibration for fingerprint maps and biased RSS measurements across devices, a gradient-based fingerprint database is designed in [17]. Meanwhile, based on the concept of the neural network, semi-supervised learning technology [18], or unsupervised learning technology [19-20] is widely applied in building a fingerprint database. In [21], a novel fingerprint database-built scheme that can adaptively adopt the proper fingerprint database according to the collected signals is proposed.

Refer to the fingerprint matching algorithm, the K-nearest neighbor (KNN) algorithm was first proposed in 2000 and its earliest corresponding localization system is the RADAR system [22], and then several optimized KNN algorithms are developed by other researchers [23-24]. For the smart building, [24] design a low-cost indoor positioning system based on the Wi-Fi fingerprint embedded on the smartphones. In this system, online layer clustering and K-nearest neighbor method based on the fisher information weighting and differential coordinates are presented, the accuracy was greatly improved. As a dimensionality reduction technique, multi-dimensional scaling (MDS) method is widely used in location awareness of indoor environments [25-26]. To obtain the input matrix between nodes, the Euclidean distance of RSS vectors between two specific positions is usually adopted to characterize their physical distance [27]. To improve the localization accuracy, literature [28] proposed a localization strategy through the collaboration of fingerprint and assistant nodes, in which the assistant nodes were selected by the similarity of RSS sequences, and the time-of-flight ranging error was mitigated by an adaptive Kalman filter. Facing with various interference factors in the indoor environments, the researchers in [29] proposed a novel scheme of fingerprint generation, representation, and matching to mitigate large errors, which yields remarkable accuracy without incurring extra cost. And in [30], a WiFi-based localization model was proposed by modifying the large localization errors and enhancing the Gaussian process regression.

## 3 Preliminary and Measurements

In this section, an RSS model at the receiver and the measures of the relationship between RSS distance vector and physical distance vector are introduced based on several given definitions.

### 3.1 Problem Statement

Nowadays, the application scenario with a large number of dense APs are very common, Figure 2 shows the RSS values collected at any position on the fifth floor of the school comprehensive building. It can be seen that the high-quality signals with RSS > - 75 are more than 10.

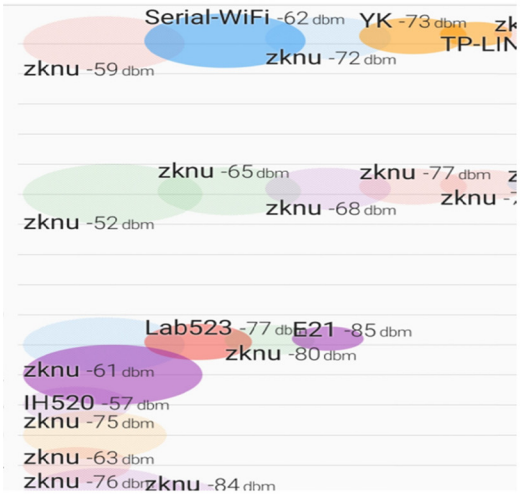
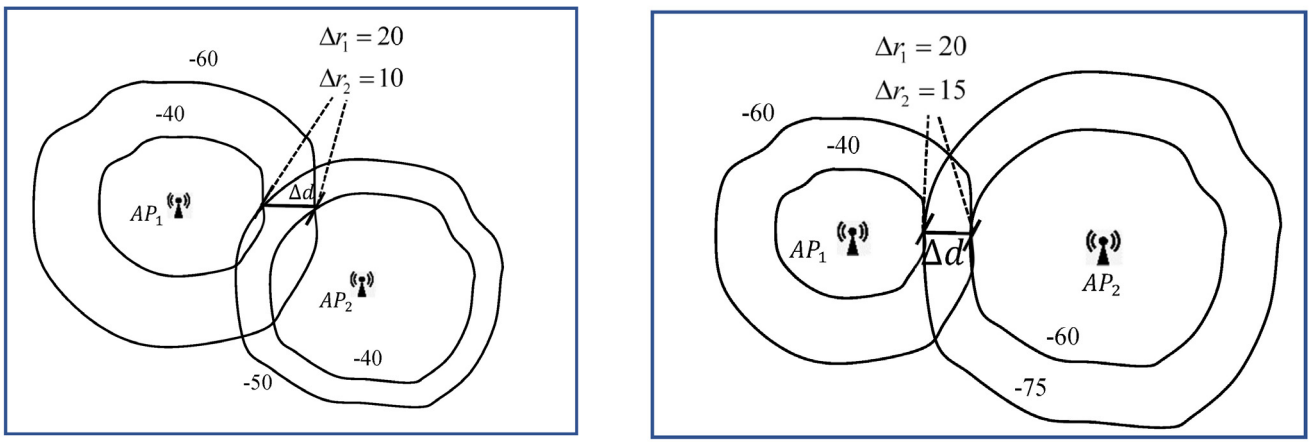



**Figure 2.** Abundant APs with high-quality signals at any Position

Subject to the propagation law of wireless signals, an identical physical distance $\Delta d$ corresponds to vastly difference of RSS change $\Delta r$. Figure 3 depicts the illustrative RSS spatial distribution of two APs, as seen, for the same $\Delta d$, $\Delta r$ is different due to the different directions (see Figure 3(a)) or the different distance (see Figure 3(b)) of the APs. Meanwhile, abnormal APs are inevitable in indoor environments. Hence, it is a very important precondition to select an appropriate and normal APs set to calculate the RSS distance for RSS distance-based localization technologies.




(a) The different direction of APs

(b) The different distance of Aps

**Figure 3.** Characterization capability diversity

This paper aims to provide a calculation algorithm of RSS distance for RSS distance-based indoor localization techniques, which is not suitable for localization techniques without RSS distance or a large number of sufficient APs scenes.

For the convenience of expression, the symbols used in the paper are shown in Table 1.

**Table 1.** The symbols used in the paper

| Symbol | | Symbol | |
|---|---|---|---|
| $m$ | The number of APs | $n$ | The number of RPs |
| $F_i$ | The fingerprint of position $i$ | $F$ | The fingerprint matrix |
| $d_{ij}^p$ | The physical distance of two positions | $d_{ij}^r$ | The RSS distance of two positions |
| $D^p$ | Physical distance vector | $D^r$ | RSS distance vector |
| $r$ | Correlation coefficient | $c$ | Consistency coefficient |
| $\Delta r$ | The RSS tree-ring distance | $\Delta d$ | The physical tree-ring distance |
| $D$ | Distance matrix $(D^P, D^r)$ | $\Delta R_j$ | RSS tree-ring distance vector |
| $E$ | The mean | $\delta$ | The standard deviation |

## 3.2 The Log-Distance Path Loss (LDPL) Model

Indoors, RSS is often affected by signal attenuation resulting from signal reflection, refraction, shielding, etc. and severe RSS fluctuation due to multipath fading and indoor noise. In particular, these phenomena are closely related to current environments. Therefore, considering the complexity of the indoor environments, the log-distance path-loss (LDPL) propagation model is commonly used to represent signal propagation loss in indoor environments, such as:

$$P_L(d) = P_L(d_0) + 10n\log(\frac{d}{d_0}) + X_\sigma \quad \textbf{(1)}$$

where $d$ is the distance between the transmitter and the receiver, measured in meter; $P_L(d)$ represents the path loss with a distance of $d$, usually measured with decibel-milliwatts (dBm); $d_0$ is the reference distance, usually 1 m, and $P_L(d_0)$ represents the path loss when the reference distance is $d_0$; $n$ is the path loss exponent, which represents the growth rate of path loss with distance and relies on the surrounding environments and building type, usually varies from 2 in free space to 4 in indoor environments. $X_\sigma$ is a Gaussian distribution random variable with a mean of zero and a standard deviation of $\sigma$. Considering environmental factors, the value of $\sigma$ generally varies between 0~14.1 dBm.

Assuming that the transmission power of the transmitter is $P_T$. From (1), the LDPL model of the signal strength at the receiver can be written as:

$$RSSI(d) = P_R(d_0) - 10n\log(\frac{d}{d_0}) + X_\sigma \quad \textbf{(2)}$$

where $P_R(D_0) = P_T - P_L(d)$. From the theoretical model in (2), for a single AP, the RSS at the receiver is closely related to the distance $d$ from the AP.

In many indoor localization techniques, RSS difference between two specific positions is often used to characterize the physical distance between them, and then the RSS difference is used to take advantage of that to further locate the positions of users or devices. However, it is well known that the RSS difference cannot characterize the physical distance well when there is only one AP. Therefore, the difference of RSS vector under multiple APs is taken to characterize the physical distance between two positions. And it is generally believed that the more APs, the better the characterization effect [27].

### 3.3 The Measure of the Relationship between RSS Distance and Physical Distance

Suppose there are $m$ APs in the positioning area, for a position $a$ with the coordinates $(x_a, y_a)$, the RSS values from $m$ APs, denoted as RSS vector $(x_{a1}, x_{a2}, ..., x_{am})$, where $x_{a1}$ is the RSS value received from the $i$th AP. Let's generate the fingerprint at position $a$ by associating the position coordinate with the RSS vector as follows:

$$F_a = (x_a, y_a, r_{a1}, r_{a2}, ..., r_{am}) \quad (3)$$

Further, the fingerprints at all $n$ positions generate a two-dimensional fingerprint matrix, as follows:

$$F = (F_1, F_2, F_n)^T \quad (4)$$

where $F_i$ represents the fingerprint at the $i$th position, and $F$ is a $n \times (m+2)$ matrix. The first two columns of $F$ are the coordinates of corresponding positions, and the last $m$ columns are RSS vectors from $m$ APs.

For the position point pair $< i, j >$ composed of any two specific position points in the positioning area, let's denote the physical distance between them as $d_{ij}^p$:

$$d_{ij}^p = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (5)$$

Since $d_{ij}^p = d_{ji}^p$, only $d_{ij}^p$, $i < j$ is considered for all positions in the positioning area. Similarly, RSS distance is defined as follows:

**Definition 1**: In a positioning area with $m$ APs, for the position point pair $< i, j >$ composed of any two specific positions, denotes the Euclidean distance between RSS vectors of position point pair $< i, j >$ as $d_{ij}^r$, $d_{ij}^r$ is called RSS distance orienting to the fingerprint of position point pair $< i, j >$. Similarly, since $d_{ij}^r = d_{ji}^r$, only $d_{ij}^r, i < j$ is considered, and $d_{ij}^r$ is calculated as follows:

$$d_{ij}^r = \sqrt{\sum_{k=1}^{m} (r_{ik} - r_{jk})^2} = \sqrt{\sum_{k=1}^{m} (10n \log \frac{d_{jk}}{d_{ik}} + \delta_k)^2} \quad (6)$$

where $d_{ik}$ and $d_{jk}$ are the distance of points $i$ and $j$ from the $k$th AP, respectively, $1 \le k \le m$.

Further, in terms of $F$, the physical distance of all position point pair composed of $n$ position points can be written in a physical distance vector as:

$$D^p = [d_{ij}^p], (i \le n, j \le n, i < j) \quad (7)$$

Similarly, the RSS distance vector is

$$D^r = [d_{ij}^r], i \le n, j \le n, i < j \quad (8)$$

In the paper, to measure the characterization capability, there are two measures for the relationship between vectors $D^p$ and $D^r$, i.e. Pearson correlation coefficient and consistency coefficient.

Statistically, the Pearson correlation coefficient is used to measure the degree of correlation (linear correlation) between two variables, which is defined as the quotient of covariance and standard deviation. When the quotient is 0.8-1.0, it is considered that the two variables are a very strong correlation. In this paper, the Pearson correlation coefficient is used to indicate the correlation between RSS distance vector and physical distance vector, to characterize the degree of linear correlation between them. It can be calculated from (9):

$$r = \frac{\sum_{i=1}^{n} \sum_{j=i+1}^{n} (d_{ij}^p - \overline{D^p})(d_{ij}^r - \overline{D^r})}{\sqrt{\sum_{i=1}^{n} \sum_{j=i+1}^{n} (d_i^p - \overline{D^p})^2} \sqrt{\sum_{i=1}^{n} \sum_{j=i+1}^{n} (d_i^r - \overline{D^r})^2}} \quad (9)$$

where $\overline{D^p}$ and $\overline{D^r}$ represent the mean value of $D^p$ and $D^r$, respectively.

In addition to the degree of linear correlation, it is usually expected that a large RSS distance corresponds to a large physical distance, which can be called the sequence consistency between two variables. In some techniques, correlations between variables are not required, as long as the sequence is consistent [27]. The consistency coefficient between $D^p$ and $D^r$ is defined as:

$$c = \frac{\Sigma_{\substack{1 \le i,j,k,l \le n \\ i<j, k<l}} f(sign(d_{ij}^r - d_{kl}^r), sign(d_{ij}^p - d_{kl}^p))}{\Sigma_{\substack{1 \le i,j,k,l \le n1 \\ i<j, k<l}}} \quad (10)$$

where $sign(x)$ and $f(x, y)$ are as follows:

$$sign(x) = \begin{cases} 1, x > 0 \\ 0, x = 0 \\ -1, x < 0 \end{cases}, f(x, y) = \begin{cases} 1, if, x = y \\ 0, \quad else \end{cases} \quad (11)$$

When the consistency coefficient $c > 0.8$, we consider RSS distance and physical distance to be very high consistent.

The above two measures are used to measure the relationship between RSS distance and physical distance in the paper.

## 4 Observation and Analysis of the Relationship between RSS Distance and Physical Distance

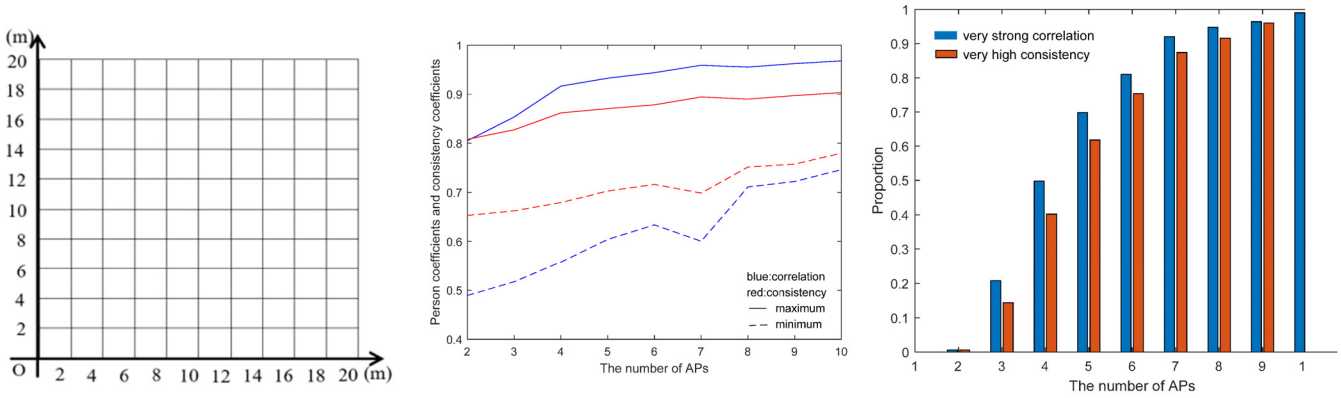
In this section, after preliminary observation, the effect of APs' positions on characterization capability is analyzed from two perspectives.

### 4.1 Experiment and Observation

Before making theoretical analysis, we conduct a simple simulation experiment to observe the effects of APs position on RSS distance of different pairwise position. For a $20m \times 20m$ indoor positioning area shown in Figure 4(a), set the interval of sampling points is $2m \times 2m$, including the points at the edges, there are 121 points where APs can be deployed. Considering the user's activity habit in practice, only 81 internal position points (excluding 40 position points at the edges) are used to sample RSS values and discuss the relationship of the pairwise position. When there is only one AP in the positioning area, there are 121 alternative positions. For an AP at each position, we calculate the measure of relationships between RSS distance and physical distance of all pairwise positions.

Among the results, the maximum and minimum Pearson correlation coefficients are $r_{\max} = 0.5091$ and $r_{\min} = 0.0242$, respectively, and the maximum and minimum consistency coefficients are $c_{\max} = 0.6547$ and $c_{\min} = 0.4314$, respectively. From the results, we can conclude that the characterization capability of RSS distance between two positions against the physical distance between them varies with the position of the AP. When there is only one AP in the positioning area, the characterization capability is very poor. Further, we study the cases of $n(n = 2...10)$ APs deployed in the positioning area. When the number of randomly deployed APs is $n(n = 2...10)$, there will be a total of $C_{121}^{n}$ alternative position combinations. For each n, randomly select 500 position combinations, recalculate the Pearson correlation coefficients and consistency coefficients between RSS distance and physical distance, and their maximum and minimum values are shown in Figure 4(b). If the APs are at the appropriate positions, even just 2 APs are enough to achieve a very strong correlation and very high consistency. However, if the APs are deployed at inappropriate positions, even 10 APs have poor correlation and consistency (lower than 0.8).

For each AP number (ranging from 2 to 10), among the 500 position combinations, the proportion of a very strong correlation and very high consistency is shown in Figure 4(c). Obviously, as the number of APs increases, so does the proportion of very strong correlation and very high consistency. However, in all deployment cases, the very strong correlation and the very high consistency between RSS distance and physical distance cannot be fully ensured (100%) even if there are 10 APs.




(a) 20m×20m indoor localization area

(b) Maximum and minimum of correlation and consistency coefficient

(c) The proportion of very strong correlation and very high consistency

**Figure 4.** Simulation results of RSS distances for different APs locations and numbers

As seen, the positions and the quantity of APs are both significant factors affecting the characterization capability of RSS distance. The latter factor has been well satisfied and the former one will be discussed in the next.

### 4.2 Physical Distance Analysis under the Identical RSS Distance

Here, suppose that for an AP at a fixed position, when the RSS distance of a position pair is known, we

discuss the corresponding physical distance.

For the convenience of analysis, with O as the center, construct the inner and outer circles with a radius of $d_1, d_2 (d_1 < d_2)$, respectively. We define the tree-ring distance as follows:

**Definition 2:** When an AP is located in the center O, A and B are respectively located on the inner and outer circumference, for the point pair $(A, B)$, the absolute value of RSS difference received from the AP is called the point pair $(A, B)$'s RSS tree-ring distance orienting to the AP, and is denoted as $\Delta r(A, B)$. Also, the radius difference of inner and outer circles is called the point pair $(A, B)$'s physical tree-ring distance and is denoted as $\Delta d(A, B)$. Obviously, the physical tree-ring distance and the RSS tree-ring distance are corresponding one by one.

**Remark 1:** RSS tree-ring distance is the RSS difference orienting to a single AP, while RSS distance is a comprehensive measure of the RSS differences of all APs in a given environment.

When there is only one AP in the positioning area, RSS tree-ring distance and RSS distance of point pairs $(A, B)$ are equal, i.e. $\Delta r(A, B) = d^r_{AB}$, and no matter how the positions of points A and B change on the inner and outer circumference, the RSS tree-ring distance and the physical tree-ring distance remain unchanged, and what changed is only the actual physical distance of point pairs $(A, B)$. Therefore, by studying the relationship between physical distance and physical tree-ring distance, we can get the change rule of physical distance under an identical RSS tree-ring distance.

Now, for an AP located at point $O$, let's consider the relationship between physical distance and physical tree-ring distance of point pair $(A, B)$. In Figure 5(a), connect $OA$, assume that the extension line of $OA$ intersects with the outer circumference at point $B'$, and the reverse extension line of $OA$ intersects with the outer circumference at point $B''$. Then, the difference between physical distance $d^p_{AB}$ and physical tree-ring distance $\Delta d(A, B)$ increases with $\theta$ ($\theta$ is the included angle between $OA$ and $OB$). Their relationship can be expressed as:

$$(d^p_{AB})^2 = d_1^2 + d_2^2 - 2d_1 d_2 con\theta \quad \textbf{(12)}$$

Rearrange (12) to obtain (13):

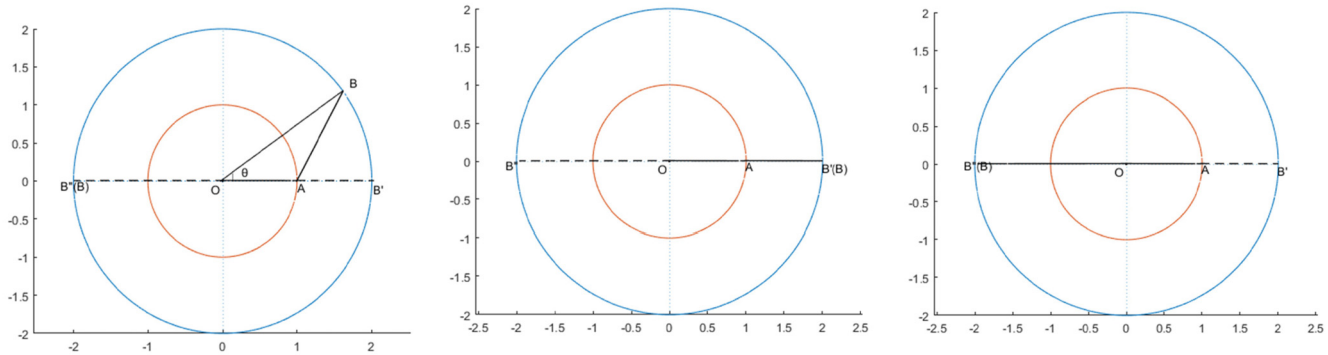
$$(d^p_{AB})^2 - (\Delta d)^2 = 2d_1 d_2 (1 - \cos\theta)(0 \le \theta \le \pi) \quad \textbf{(13)}$$

Then, we draw the following conclusions from (13):

(1) The quadratic difference of physical distance $d^p_{AB}$ and physical tree-ring distance $\Delta d$ of point pair $(A, B)$ monotonically increase with $\theta$;

(2) As shown in Figure 5(b), when $\theta = 0$, point $B$ coincides with point $B'$, the physical distance $d^p_{AB}$ is equal to physical tree-ring distance $\Delta d$, i.e. $d^p_{AB} = \Delta d = d_2 - d_1$, the AP is just on the reverse extension line of AB. In this case, the position relationship between $(A, B)$ and the AP is optimal, and RSS distance between them is the most accurate characterization to the physical distance;

(3) As shown in Figure 5(c), when $\theta = \pi$, point $B$ coincides with point $B''$, and the quadratic difference of physical distance $d^p_{AB}$ and physical tree-ring distance $\Delta d$ is the maximum, i.e. $d^p_{AB} - \Delta d = 2d_1$, and the AP is just on line AB. In this case, the characterization capability of RSS distance against physical distance $(A, B)$ is the poorest.




(a) The common case (b) The best case (c) The worst case

**Figure 5.** The position change case of point pair $(A, B)$ under the identical RSS distance

Note that, when $d_1$ is very small, that is, point $A$ is close enough to the AP, the physical distance $d^p_{AB}$ for any point $A$ and point $B$ is approximately equal to physical tree-ring distance $\Delta d(A, B)$, then, the RSS distance can characterize all physical distance $(A, B)$ well.

In sum, for an AP in the positioning area, the physical distance characterized by the identical RSS distance is not equal except there exists an AP near one point of the point pair.

### 4.3 RSS Distance Analysis under the Identical Physical Distance

Now, from another perspective, assume that the physical distance of $(A, B)$ is determined, i.e. $d^p_{AB} = d_0$, we discuss the changes of RSS distance generated by APs at different positions.

Suppose the position point pair $(A, B)$ is fixed, an AP at point $O$. With $O$ as the center, construct inner and outer circles with the radius of $d_{OA}$ and $d_{OB}$, respectively. Connect $OB$ to intersect with the inner circumference at point $A'$, connect $OA$ to intersect with the outer circumference at point $B'$, as shown in Figure 6(a). From the foregoing discussion, $d^p_{AB} = d^p_{AB'} = \Delta d$, that is, $d^p_{A'B}$ is the physical tree-ring distance of position point pair $(A, B)$. As discussed above, by studying the changes of $\Delta d(A, B)$ orienting different positions of APs, we can obtain the changes of RSS distance.

For $\Delta AA'B$, the following equation holds:

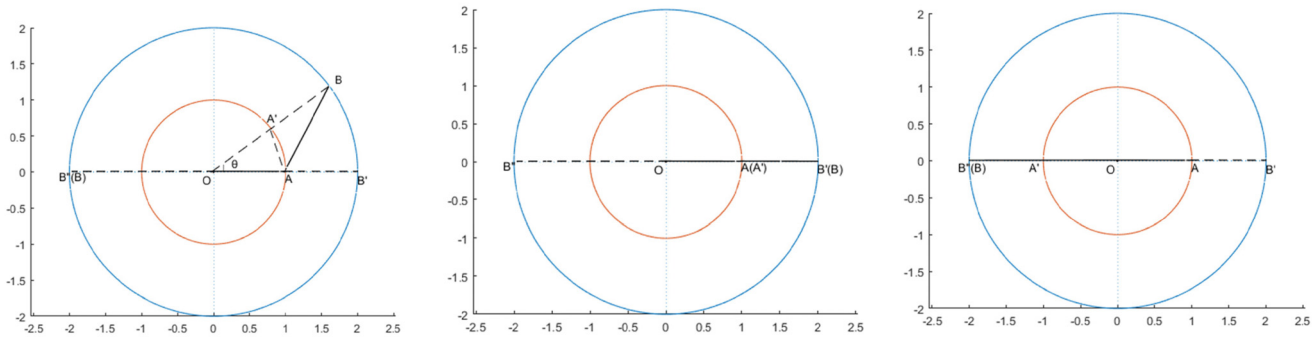
$$\Delta(A, B) = d^p_{A'B} = d_0 - 2d_{OA} - \sin\frac{\theta}{2} \tag{14}$$

From (14), we can conclude that:

(1) The physical tree-ring distance $\Delta d(A, B)$ monotonically decreases with $\theta(0 \le \theta \le \pi)$, and the smaller $\theta$ is, the bigger the physical tree-ring distance $\Delta d(A, B)$ is and the closer it get to $d_0$;

(2) When $\theta = 0$, the AP is on the extension line of $BA$, point $A'$ coincides with point $A$ and point $B'$ coincides with point $B$, $\Delta d = d_0$. In this case, RSS distance of position point pair $(A, B)$ has the optimal characterization capability against physical distance (Figure 6(b)).

(3) When $\theta = \pi$, the AP is located on the line of $AB$. In this case, $\Delta d = d_0 - 2d_{OA}$, the characterization capability of RSS distance against physical distance is the poorest (Figure 6(c)).




(a) The Common Case (b) The Best Case (c) The Worst Case

**Figure 6.** The position change case of point pair $(A, B)$ for the determined physical distance

Note that when $d_{OA}$ is very small, that is, the AP is close enough to point $A$, the physical tree-ring distance $\Delta d(A, B)$ is approximately equal to physical distance $d_0$, then, the RSS distance generated by any AP close enough to $A$ can characterize the physical distance $(A, B)$ well. In sum, if an AP at different positions of the positioning area, the RSS distance is not equal for the same position pair except the AP is near one position of the position pair.

## 5 RSS Distance Calculation Algorithm Based on Tree-ring Distance References

In this section, the way of APs deployment and selection are first summarized, and then an algorithm for calculating the RSS distance of each position pair is presented.

### 5.1 Deployment and Selection of APs in Different Application Scenarios

For a position point pair$(A, B)$ with known physical distance and an AP at position point O, suppose that the distance between the AP and position point A is $d_{OA}$ and $d_{OA} < d^p_{OA}$, construct a circle with point $A$ as the center and $d_{OA}$ as the radius, as shown in Figure 7. So, for an AP at any position on the circumference, the RSS value collected at position A is identical in theory, but the RSS value at position B is diversity. As a result, for an AP at different positions, the RSS distance of the same position point pair $(A, B)$ are different. According to the foregoing analysis, the optimal and poorest characterization capability of RSS distance to physical distance corresponds to that AP at point $O'$ and point $O''$, respectively. Hence, we can conclude that when an AP is located on or near the line of the position

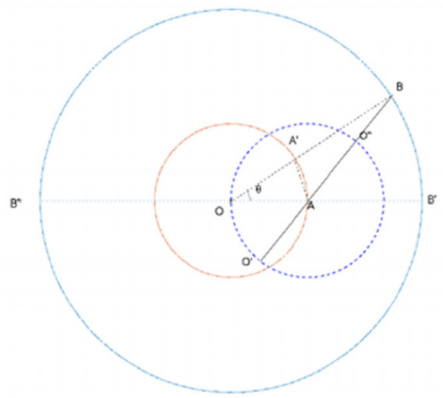
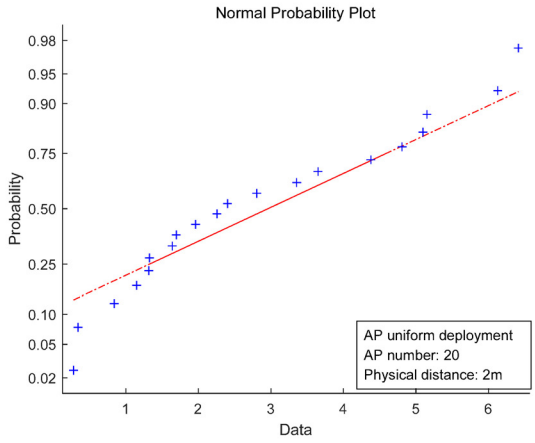


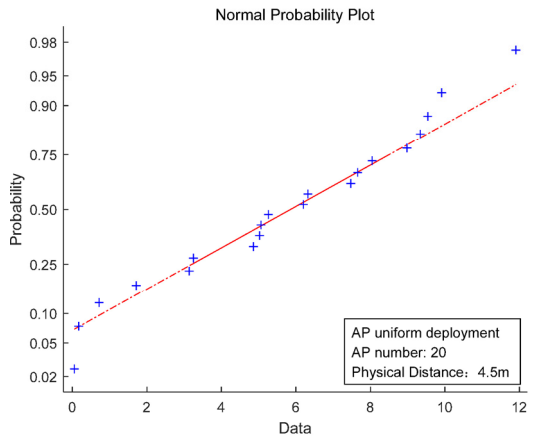

**Figure 7.** The effect of APs position on RSS distance for the fixed-point pair ($A$, $B$)

point pair ($A$, $B$), the corresponding RSS distance has the poor characterization capability to the physical distance, that is, the RSS distance is small, while when an AP is located on the same side of line $AB$, the corresponding RSS distance has good characterization capability, that is, the RSS distance is large. However, near the line of point pair ($A$, $B$) is exactly the region where APs is deployed to facilitate communication. Therefore, for the position point pair ($A$, $B$), we must select the "right" APs set with suitable positions from the rich APs to calculate the RSS distance.

Overall, for a position pair, it is more appropriate to select APs who make the tree-ring distance of the position pair larger. The above analysis shows that these APs should be on the same side of the pairwise position and the smaller the deviation from the extension line connecting of the two positions, the better it is. Also, we should exclude the large RSS tree-ring distance caused by the abnormal APs. For any pairwise position, all RSS tree-ring distance orienting to all APs are normally distributed. To make the results clearer and easier to observe, we have deployed 20 APs in the positioning area, Figure 8 shows the RSS tree-ring distance distribution when the physical distance is 2 meters and 4.5 meters respectively. Hence, it is reasonable to use Means+3S to eliminate abnormal APs.




(a) The physical distance is 2m




(b) The physical distance is 4.5m

**Figure 8.** The RSS tree-ring distance distribution

Here, we summarize the guidance for APs deployment and selection as follows:

(1) For some special application scenarios, such as telecare and telehealth, APs are installed mainly for positioning, the edge triangle deployment makes the overall effect better.

(2) For the indoor environments where many APs already exist, the guidance for selecting APs is to select the APs that make the RSS tree-ring distance of a position pair larger.

## 5.2 Tree-ring Distance Based RSS Distance Calculation Algorithm

Based on the foregoing analysis, we know that when an AP is located on or near the line of a position pair, the RSS tree-ring distance is small (because the physical tree-ring distance corresponds to the RSS tree-ring distance, in this section, RSS tree-ring distance is used for discussion), that is, the RSS tree-ring distance of the AP has a weak ability to characterize its physical distance, so the AP is not recommended for participating in the calculation of RSS distance. Besides, to avoid large RSS tree-ring distance caused by abnormal APs in the complex indoor environments, the method of Mean+3S is used to eliminate them.

For an indoor positioning area with $m$ APs and $n$ RPs, the distance matrix of all RPs' pairwise positions is written as $D=(D^p, D^r)$, where $D^p$ is the physical distance vector as defined in (7) and $D^r$ is the RSS distance vector as defined in (8) . So, the element of D is as follows:

$$D_j=(D^p, D^r)=(D^p_{ii'}, D^r_{ii'})(1\le j\le C_n^2, i, i') \quad (15)$$

where $d_{ii'}^{p}$ and $d_{ii'}^{r}$ represent the physical distance and RSS distance of position point pair $(i,i')$ of the $i$th and the $i'$ th RP, respectively, and the relationship between $i$ and $j$ satisfies $j=\sum_{t=1}^{i-1}(n-t)+(i'-1)$ 。

To got $D^r$ , the APs involved in calculating $d_{ii'}^{r}$ according to (8) need to be determined first. Now, by comparing the difference of RSS tree-ring distance orienting to each APs, we present an algorithm to select the right APs. The calculation process of $d_{ii'}^{r}$ is exactly the process in which APs compete to participate in calculating RSS distance.

Firstly, RSS tree-ring distance vector of point pair $(i,i')$ orienting to all APs is denoted as follows:

$$\Delta R_j=(\Delta r_1,\Delta r_2,\Delta r_m)=(1\le j\le C_n^2,i,i') \tag{16}$$

where $\Delta n_k(1<k<m)$ represents the RSS tree-ring distance of position point pair $(i,i')$ orienting to the $k$th AP.

Then, the elements in vector $\Delta R_j$ are sorted in descending order, and the sorted vector is denoted as:

$$\Delta\tilde{R}_j=(\Delta\tilde{r}_1,\Delta\tilde{r}_2,\Delta\tilde{r}_m) \tag{17}$$

Further, the mean and standard deviation of $\Delta\tilde{R}_j$ are calculated by (18) and (19), respectively:

$$E(\Delta\tilde{R}_j)=\sum_{k=1}^{m}\Delta\tilde{r}_k \tag{18}$$

$$\delta(\Delta\tilde{R}_j)=(\frac{1}{m}\sum_{k=1}^{m}(\Delta\tilde{r}_k-e(\Delta\tilde{R}_j))^2)^{\frac{1}{2}} \tag{19}$$

Next, for $\Delta\tilde{R}_j$ , the difference between the header element and the mean value is calculated to determine whether it exceeds three times the standard deviation. If it is, the corresponding APs will be eliminated. Then, again, calculate the mean, the standard deviation, the difference between the header element and the mean of remaining RSS tree-ring distance vector, and judge it. So repeatedly, until their difference does not exceed three times the standard deviation.

Finally, for the final vector $\Delta\tilde{R}_j$ , select the first $t_0$ values, and calculate RSS distance $d_{ii'}^{r}$ of position pair $(i,i')$ according to formula (6). For a position pair $D_j^r$ , the flow chart of the RSS distance calculation algorithm is described in Figure 9.

Similarly, we can obtain the RSS distance of all RPs' pairwise position.

**Remark 2:** As for $t_0$, it depends on the actual environments. In the common indoor environments of section 5, $t_0=3$, which meets the requirements of correlation and consistency between RSS distance and physical distance.

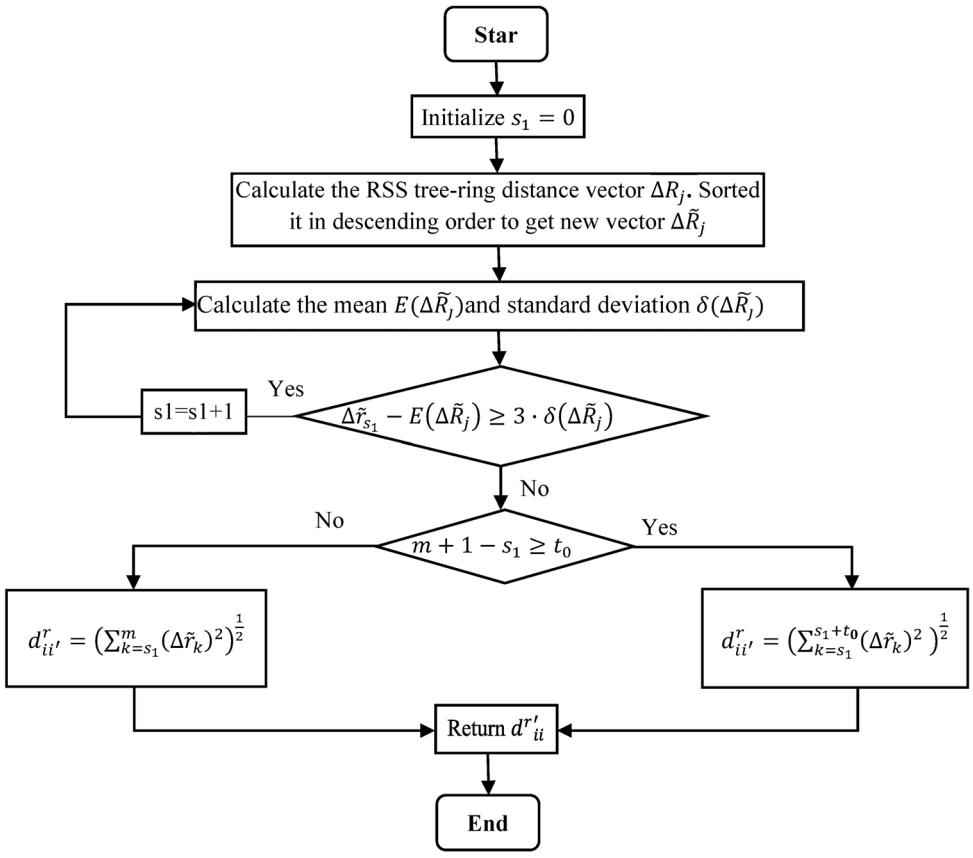



**Figure 9.** The flow chart of the RSS distance calculation algorithm

**Remark 3:** Some references distinguish the degree of participation of APs by setting the weights of APs [8, 27]. Unlike these studies, we believe that the same APs have different effects on different position pairs. Therefore, it makes sense to identify different APs set for the different pairwise position.

## 6 Simulations and Experiment

Here, we conduct some simulation experiments to evaluate the algorithm. The experimental parameters and presets are described in Table 2.

**Table 2.** The parameters and presets of the experiments

| Simulation Platform | MATLAB |
|---|---|
| Positioning Area | $20\times 20 = 400\ m^2$ |
| Interval of APs | $2m\times 2m$ |
| Propagation Model | LDPL |
| The Number of APs | 10, randomly deployed, 57 alternative positions (red dots in Figure 10) |
| The Number of RPs | 81, internal position points (excluding 40 position points at the edges) |

Note that, to simulate the real indoor environments, we calculate multiple propagation paths between the transmitter (AP) and the receiver (smart device at an RP), including one direct path and six wall reflection paths. Considering the influence of multiple reflections is small, it is ignored.

We divided the simulated indoor environment into 100 grids. There is a total of 121 cross points, see Figure 10. 81 internal position points are used as the reference points, which form $C_{91}^{2} = 3240$ position point pairs.

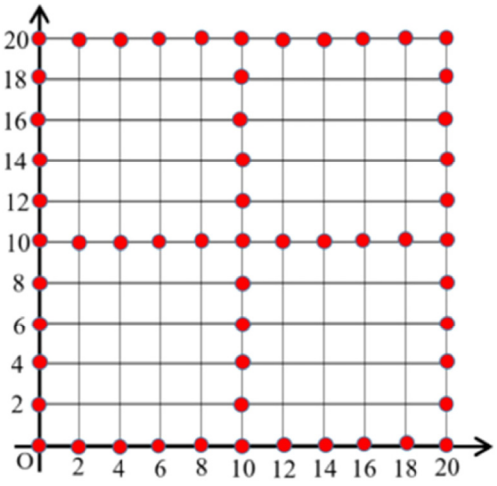



**Figure 10.** The alternative deployment positions for APs

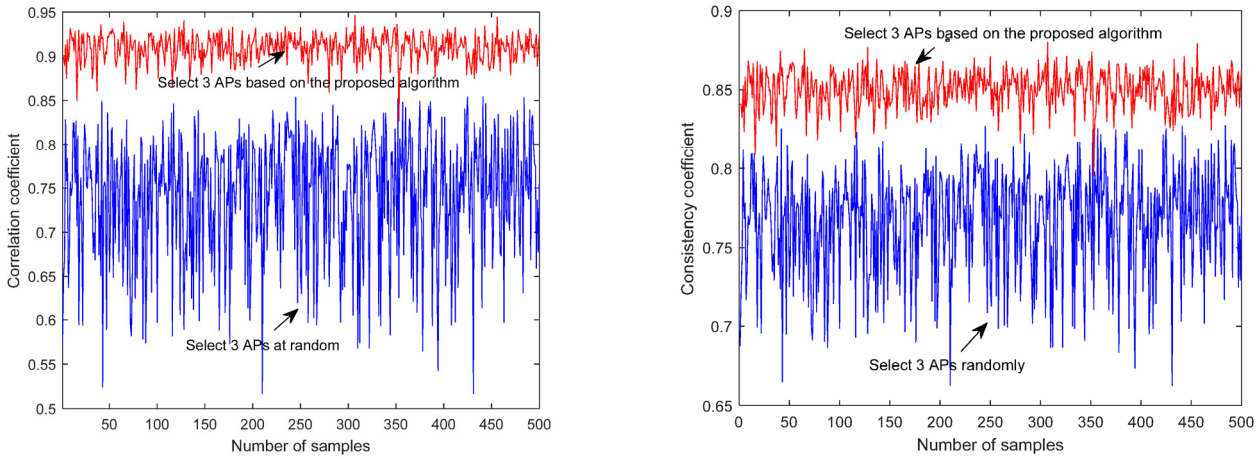
Due to the autonomy and flexibility of APs deployment in public indoor environments, there are $C_{57}^{10}$ kinds of APs deployments combinations theoretically. We take 500 of them as randomly selected APs position combinations. Note that it is reasonable to have about 5 APs in 400 $m^2$ indoor environments.

In the realistic scenario, it is reasonable to have 5 APs in the $400m^2$ However, in our simulation, when 5 APs are randomly deployed in 57 positions for 500 times, it is easy to generate the uneven distribution of APs, which is inconsistent with the actual situation. Moreover, in the case of ensuring the uniform distribution of APs, we only select three of them to participate in the calculation of RSS distance. Therefore, redundant APs do not affect the simulation effect.

For each APs position combinations, we calculate the RSS distance of all pairwise positions in two cases, that is, (1) select 3 APs from 10 APs randomly, and (2) select 3 APs according to the algorithm proposed in this paper. Then, the 500-correlation coefficient and the consistency coefficient are shown in Figure 11. As seen, in 500 cases, using the proposed algorithm, the correlation coefficient and consistency coefficient between RSS distance and physical distance are both greater than 0.8, that is, RSS distance and physical distance are very strong correlation and very high consistent. In contrast, when the 3 APs are randomly selected, only a small part of the 500 cases, RSS distance and physical distance are very strong correlation and very high consistent. Through the statistics of the results, the proportion of very strong correlation and very high consistency in 500 cases is only 20.8% and 14.4% respectively.

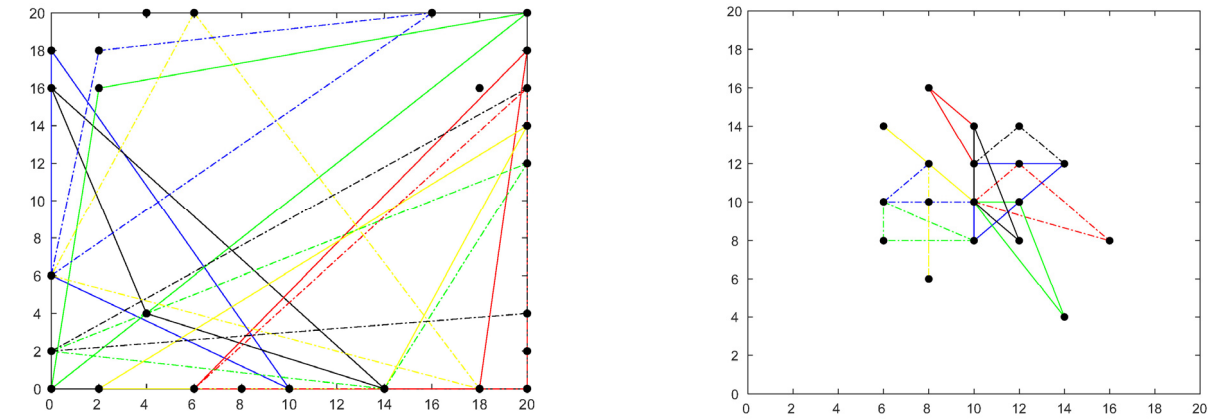
Since the efficiency and quality of the RSS distance provided, the higher the efficiency and accuracy of the corresponding localization algorithm. Considering the complexity and APs numbers of the indoor environments, we suggest that at least 3 APs are selected for RSS distance calculation in the actual localization process, i.e. $t_0 \geq 3$ in the algorithm process.

Furthermore, to demonstrate the edge deployment of APs makes the overall effect better, for the experiments in section 4.1, we repeat the simulation of the number of APs $n = 3$ for 10 times. For each simulation, we record the positions of APs set when the relationship coefficient is maximum and minimum in 500 repeated experiments. We find that the APs positions of the maximum correlation coefficient and the maximum consistency coefficient are the same, so is the case of the minimum value. The 10 groups of positions are shown in Figure 12. Obviously, for the combination of the optimal positions, all positions are evenly distributed in the edge positions (see Figure 12(a)), while for the combination of poor positions, they are mostly concentrated in the center region of the positioning area (see Figure 12(b)).




(a) Comparison of the correlation coefficient (b) Comparison of consistency coefficient

**Figure 11.** The Comparison of the Proposed Algorithm-based Selection and Random Selection for 3 APs



(a) The optimal positions combination of 3 APs (b) The poor position combinations of 3 APs

**Figure 12.** The positions of APs subset

## 7 Conclusions

In this paper, we propose an algorithm regarding the improvement and optimization of RSS distance calculation, which can improve the key steps of all localization technologies based on RSS distance. Next, we will verify the improvement effect of the proposed algorithm in the corresponding RSS distance based localization techniques. Also, the future work includes (1) The improved fingerprint collection and updating algorithms to reduce work intensity; (2) The improvement of location accuracy with additional auxiliary equipment, such as Bluetooth, RFID, infrared and various sensors and so on; (3) The customized application services with different location accuracy according to different scene requirements.

## Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant 61702071 and 61501076, High-Level Talents Scientific Research Foundation of Zhoukou Normal University under grant ZKNUC2017024, Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis (Open-ended fund under grant GDUPTKLAB201505) and Key Scientific research projects of colleges and universities in Henan Province under grant 19B520031.

## References

[1] F. Zafari, A. Gkelias, K. K. Leung, A Survey of Indoor Localization Systems and Technologies, *IEEE Communications Surveys & Tutorials*, Vol. 21, No. 3, pp. 2568-2599, Third Quarter, 2019.

[2] H. J. Huang, J. G. Zhou, W. Li, J. B. Zhang, X. Zhang, G. L. Hou, Wearable Indoor Localisation Approach in Internet of Things, *IET Networks*, Vol. 5, No. 5, pp. 122-126, September, 2016.

[3] C. -H. Cheng, Y. Yan, Y. -F. Huang, Performance of Improved Fuzzy Indoor Zone Positioning Systems in Wireless Sensor Networks, *Journal of Internet Technology*, Vol. 19, No. 4, pp. 1241-1250, July, 2018.

[4] F. Liu, J. Liu, Y. Q. Yin, W. H. Wang, D. H. Hu, P. P. Chen, Q. Niu, Survey on WiFi-based Indoor Positioning Techniques, *IET Communications*, Vol. 14, No. 9, pp. 1372-1383, June, 2020.

[5] W. Huang, Y. Xiong, X. Li, H. Lin, X. Mao, P. Yang, Y. Liu, X. Wang, Swadloon: Direction Finding and Indoor Localization Using Acoustic Signal by Shaking Smartphones, *IEEE Transactions on Mobile Computing*, Vol. 14, No. 10, pp. 2145-2157, October, 2015.

[6] Q. Gao, C. Shen, X. Chen, K. Zhang, Weighted Moving Average-based Differential UWB Indoor Localisation System for High External Disturbance Environment, *International Journal of Internet Protocol Technology*, Vol. 12, No. 3, pp. 136-143, July, 2019.

[7] M. Hasani, J. Talvitie, L. Sydänheimo, E. Lohan, L. Ukkonen, Hybrid WLAN-RFID Indoor Localization Solution Utilizing Textile Tag, *IEEE Antennas and Wireless Propagation Letters*, Vol. 14, pp. 1358-1361, February, 2015.

[8] J. Luo, X. Yin, Y. Zheng, C. Wang, Secure Indoor Localization Based on Extracting Trusted Fingerprint, *Sensors*, Vol. 18, No. 2, pp. 469, February, 2018.

[9] A. Basiri, E. S. Lohan, T. Moore, A. Winstanley, P. Peltola, C. Hill, P. Amirian, P. Figueiredo e Silvab, Indoor location Based Services Challenges, Requirements and Usability of Current Solutions, *Computer Science Review*, Vol. 24, pp.1-12, May, 2017.

[10] Z. Yang, Z. Zhou, Y. Liu, From RSSI to CSI: Indoor Localization via Channel Response, *ACM Computing Surveys (CSUR)*, Vol. 46, No. 2, pp. 1-32, November, 2013.

[11] D. B. Ninh, J. He, V. T. Trung, D. P. Huy, An Effective Random Statistical Method for Indoor Positioning System Using WiFi Fingerprinting, *Future Generation Computer Systems*, Vol. 109, pp. 238-248, August, 2020.

[12] C. Gentile, N. Alsindi, R. Raulefs, C Teolis, *Geolocation Techniques: Principles and Applications*, Berlin, Germany, Springer Science & Business Media, 2012.

[13] J. S. Leu, M. C. Yu, H. J. Tzeng, Improving Indoor Positioning Precision by Using Received Signal Strength Fingerprint and Footprint Based on Weighted Ambient Wi-Fi Signals, *Computer Networks*, Vol. 91, No. 1, pp. 329-340, November, 2015.

[14] C. Wu, Z. Yang, Y. Liu, Smartphones Based Crowdsourcing for Indoor Localization, *IEEE Transactions on Mobile Computing*, Vol. 14, No. 2, pp. 444-457, February, 2015.

[15] S. He, W. Lin, S.-H. G. Chan, Indoor Localization and Automatic Fingerprint Update with Altered AP Signals, *IEEE Transactions on Mobile Computing*, Vol. 16, No. 7, pp. 1897-1910, July, 2017.

[16] C. Song, J. Wang, WLAN Fingerprint Indoor Positioning Strategy Based on Implicit Crowdsourcing and Semi-Supervised Learning, *International Journal of Geo-Information*, Vol. 6, No. 11, pp. 356, November, 2017.

[17] Y. Shu, Y. Huang, J. Zhang, P, Coué, P. Cheng, J. M. Chen, K. G. Shin, Gradient-based Fingerprinting for Indoor Localization and Tracking, *IEEE Transactions on Industrial Electronics*, Vol. 63, No. 4, pp. 2424-2433, April, 2016.

[18] M. Mohammadi, A. Al-Fuqaha, Enabling Cognitive Smart Cities Using Big Data and Machine Learning: Approaches and Challenges, *IEEE Communications Magazine*, Vol. 56, No. 2, pp. 94-101, February, 2018.

[19] S. Jung, B. Moon, D. Han, Unsupervised Learning for Crowdsourced Indoor Localization in Wireless Networks, *IEEE Transactions on Mobile Computing*, Vol. 15, No. 11, pp. 2892-2906, November, 2016.

[20] J. H. Seong, D. H. Seo, Selective Unsupervised Learning-Based Wi-Fi Fingerprint System Using Autoencoder and GAN, *IEEE Internet of Things Journal*, Vol. 7, No. 3, pp. 1898-1909, March, 2020.

[21] C. -S. Hsu, Y. -S. Chen, T. -Y. Juang, Y. -T. Wu, An Adaptive Wi-Fi Indoor Localisation Scheme Using Deep Learning, *International Journal of Ad Hoc and Ubiquitous Computing*, Vol. 30, No. 4, pp. 265-274, April, 2019.

[22] P. Bahl, V. N. Padmanabhan, RADAR: An In-building RF-based User Location and Tracking System, *IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No. 00CH37064)*, Tel Aviv, Israel, Vol. 2, pp. 775-784.

[23] C. -Y. Yang, Y. -W. Ma, J. -L. Chen, C. -J. Lin, W. -L. Lee, Novel Dynamic KNN with Adaptive Weighting Mechanism for Beacon-based Indoor Positioning System, *Journal of Internet Technology*, Vol. 20, No. 5, pp. 1601-1610, September, 2019.

[24] M. B. Afuosi, M. R. Zoghi, Indoor Positioning Based on Improved Weighted KNN for Energy Management in Smart Buildings, *Energy and Buildings*, Vol. 212, 109754, April, 2020.

[25] N. Saeed, H. Nam, T. Y. Al-Naffouri, M.-S. Alouini, A State-of-the-Art Survey on Multidimensional Scaling-Based Localization Techniques, *IEEE Communications Surveys & Tutorials*, Vol. 21, No. 4, pp. 3565-3583, Fourth Quarter, 2019.

[26] W. Cui, C. Wu, W. Meng, B. Li, Y. Zhang, L. Xie, Dynamic Multidimensional Scaling Algorithm for 3-D Mobile Localization, *IEEE Transactions on Instrumentation and Measurement*, Vol. 65, No. 12, pp. 2853-2865, December, 2016.

[27] Y. L. Xiao, S. G. Zhang, J. X. Wang, An Indoor Localization Algorithm Based on Multidimensional Scaling and Region Refinement, *Chinese Journal of Computers*, Vol. 40, No. 8, pp. 1918-1932, August, 2017.

[28] Q. Li, W. Li, W. Sun, J. Li, Z. Liu, Fingerprint and Assistant Nodes Based Wi-Fi Localization in Complex Indoor Environment, *IEEE Access*, Vol. 4, pp. 2993-3004, June,

2016.

[29] C. S. Wu, Z. Yang, Z. M. Zhou, Y. H. Liu, M. Y. Liu, Mitigating Large Errors in WiFi-based Indoor Localization for Smartphones, *IEEE Transactions on Vehicular Technology*, Vol. 66, No. 7, pp. 6246-6257, July, 2017.

[30] M. Xue, W. Sun, H. S. Yu, H. W. Tang, A. P. Lin, X. Zhang, R. Zimmermann, Locate the Mobile Device by Enhancing the WiFi-based Indoor Localization Model, *IEEE Internet of Things Journal*, Vol. 6, No. 5, pp. 8792-8803, October, 2019.

## Biographies

**Baofeng Wang** received the M.S. degree and Ph.D. degrees from Dalian Maritime University, Dalian, in 2006 and 2010, respectively. After graduation, she worked in Dalian University, and joined Zhoukou Normal University, China, in June 2017. Her research interests include Smart Health and Medical, Indoor Localization and State Estimation.

**Liming Chen** (M'10, SM'20) received his B.E. and M.E. degrees from the Beijing Institute of Technology, Beijing, China, and the Ph.D. degree from De Montfort University, Leicester, U.K. He is currently a professor of University of Ulster. His research interests include data analytics, pervasive computing, artificial intelligence, activity recognition, smart environment, digital healthcare and assisted living.

**Zumin Wang** Ph.D., professor in College of Information Engineering, Dalian University, China, research interests mainly focused on the theory and technologies of the Internet of Things.

**Mengmeng Xu** received the M.S. degree from Zhengzhou University, China, in 2012, and the Ph.D. degree in communication and information systems from Xidian University, China, in 2016. Since 2017, he has been with Zhoukou Normal University, China. His research interests include wireless sensor networks and cloud computing.

**Shuai Tao** received the M.S. degree from China Agricultural University, China in 2009, and the Ph.D. degree from Hokkaido University, Japan, in 2013. He is an Associate Professor of Dalian University, China. His current research interests include ambient intelligence, recognition of activities of daily living, smart medical and gait analysis.