A WKNN Indoor Fingerprint Localization Technique Based on Improved FD's Discrimination Capability to PD

Baofeng Wang , Qinghai Li , Jia Liu * , Zumin Wang , Qiudong Yu , Rui Liang

Posted Date: 15 January 2024

doi: 10.20944/preprints202401.1075.v1

Keywords: indoor fingerprint localization; WKNN; fingerprint distances

Article

A WKNN Indoor Fingerprint Localization Technique Based on Improved FD's Discrimination Capability to PD †

Baofeng Wang [1], Qinghai Li [1], Jia liu [1,*], Zumin Wang [2], Qiudong Yu [1] and Rui Liang [1]

[1] School of Information Technology Engineering, Tianjin University of Technology and Education, Tianjin, 300222, China; wangbaofeng@tute.edu.cn (B. W.); 0303180108@tute.edu.cn (Q. L); lrcalr@163.com (R.L.)
[2] College of Information Engineering, Dalian University, Dalian, Liaoning, 116600, China; wangzumin@163.com
* Correspondence: liujia78928@126.com
† This paper is an extension of work originally presented in IEEE SmartIoT 2023.

Abstract: The weighted K-nearest neighbors (WKNN) algorithm has been widely used in indoor fingerprint localization, which utilizes fingerprint distance (FD) of pairwise positions to discriminate the physical distance (PD) between them. However, due to the varied states of different wireless access points (APs), their contribution to FD discrimination capability varies. In our study, we analyzed several critical causes that affect APs' contribution, including AP's health state, the distance, and the direction between APs and position pairs. Inspired by these insights, a threshold was set for all sample RSS to eliminate the impact of abnormal APs on FD, and a correction quantity for each RSS was provided by the distance between APs and sample positions to reduce the strong signal influence on FD. Furthermore, a priority weight was designed by RSS differences (RSSD) to further optimize FD's capability to discriminate PD. Integrating the above policies, a new indoor fingerprint localization technique is redefined, referred to FD's discrimination capability improvement WKNN (FDDC-WKNN), which is suitable for indoor scenes with a large number of APs that are not uniformly managed. Our simulation results show that the correlation and consistency between FD and PD are well improved, with strong correlation increasing from 0 to 76% and high consistency increasing from 26% to 99%, which confirms that the proposed policies can greatly enhance FD's discrimination capabilities to PD, and we also found that abnormal APs can cause significant impact on FDs' discrimination Capability. Further, by implementing the FDDC-WKNN algorithm in experiments, we obtained the optimal K value in both the simulation scene and real library scene, under which the average localization errors have been reduced from $2.2732m$ to $1.2183m$ and from $3.4295m$ to $2.2068m$, respectively. In addition, compared to not using the FDDC-WKNN, the cumulative distribution function (CDF) of the localization errors curve converged faster and the errors fluctuation is smaller, which demonstrate FDDC-WKNN having stronger robustness and more stable localization performance to the state of APs in indoor environments.

Keywords: indoor fingerprint localization; WKNN; fingerprint distances

1. Introduction

With the rapid development of wireless communication technology in pervasive infrastructure and mobile clients, the demand for LBSs (location-based services) is also accelerating in people's daily lives. Hence, indoor localization techniques have got great concern in both academia and industry, and various indoor localization techniques have been developed, classified by the dependent infrastructure, such as Wi-Fi [1] [2], acoustic signals [3], UWB [4], RFID [5], etc.; by the nature of signals to be measured, such as TDOA(Time Difference Of Arrival), TOA (Time Of Arrival), RSS (Received Signal Strength) [6], and CSI (Channel State Information)[7], etc.

In particular, because of the advantages of no need for additional infrastructure, low cost, and high localization accuracy, especially with the rapid development and popularization of smartphone and WLAN infrastructure, indoor fingerprint localization has become one of the focuses of current research. Among all developed indoor localization techniques, the WLAN-based localization method

is one of the most popular methods. Which, indoor fingerprint localization is widely popular because it does not require the accurate location of the APs and can fully utilize public WLAN facilities for localization [8]. The fingerprint-based indoor localization is usually conducted in two phases, i.e. offline site survey phase and online fingerprint matching phase. Accuracy has long been the primary challenge for indoor fingerprint localization. So far, a lot of study work and achievements have been made in these two phases and many systems, and techniques and technologies have been developed to reduce localization error and improve localization accuracy [9] [10].

The multidimensional scaling (MDS) technology is a data dimensionality reduction method widely applied in indoor localization [11], which employs FDs to represent PDs to construct a distance matrix between nodes to obtain their similarity or correlation, thereby calculating an estimate of TP [12]. In recent years, with the vigorous development of artificial intelligence (AI) technology, a large number of machine learning have been introduced into indoor fingerprint localization, such as support vector machine [13], weighted K-nearest neighbor (WKNN) [14] [15], and artificial neural network (ANN) [16] [17], and so on. To deal with high noise in RSSI measurements and ensure high target-localization accuracy, authors in [13] proposed two range-free target-localization schemes on SVR, one is a plain support vector regression (SVR)-based model and the other is a fusion of SVR and Kalman filter (KF). WKNN has been widely used in indoor localization, for the problem of traditional WKNN not getting high accuracy and having poor stability, the Manhattan distance was introduced in [14], while in [15], a combined algorithm based on WKNN and extreme gradient boosting (XGBoost) was proposed. Further, the BP (Back Propagation) neural network is used in [17] to determine the distances between the TP and each RP to mitigate the effect of fluctuation of RSS. And deep neural network (DNN) model is considered in [18], where DNN models obtained from multiple training sessions are combined to locate the target, achieving the lower RMSE. In all, it has been proved that machine learning can effectively improve indoor localization accuracy, enhance system robustness, reduce costs, and improve the performance of the indoor localization method [19].

In addition, a variety of other localization strategies have also been proposed. For example, in literature [20], a location-aware infrastructure is presented, which can give rise to a sensing mechanism that enables infacility crowdsourcing can help fingerprinting localization services. By the similarity of RSS sequences, the assistant nodes were selected in [21].to improve the localization accuracy, and by an adaptive Kalman filter, the time-of-flight ranging error has been alleviated. Recently, a fingerprint augment framework based on super-resolution (FASR) was proposed in [22] to reduce the cost at the offline phase and ensure localization accuracy. Based on the conversion between fingerprint image and fingerprint data, the framework achieved the fusion of fingerprint augment and super-resolution. A transportable laser range scanner was used to automatically label Wi-Fi scans in [23] to get an accurate indoor fingerprint localization system with no associated data collection.

For the online phase of fingerprint-based localization, it is widely known that fingerprint matching techniques play an important role in reducing localization errors, and have been developed various techniques [24] [25]. It can be seen that many fingerprint-based localization algorithms are based on a common premise, that is to use FD to discriminate PD, while FD is the Euclidean distance of RSSD [26] [27]. In [28], we have discussed that RSSD represents the difference in RSS received by an AP at two different locations, which is closely related to the state of the APs. Nowadays, with the booming development of intelligence, many indoor spaces can receive a large number of AP signals. These APs with high-quality signals are installed for communication, they are evenly distributed and dense, but they are not uniformly managed. That is to say, we do not know the state of most of APs. When we utilize these ubiquitous APs for localization, they will bring greater interference and uncertainty, seriously affecting localization performance.

However, the root cause of localization errors from APs has not been adequately studied. Although researchers in [29] identified several crucial problems that caused the localization errors, only the different distances of the APs having diverse discrimination capabilities were studied. In [30], the problem of the impaired APs and the joining of the new APs were considered and a secure

 doi:10.20944/preprints202401.1075.v1

fingerprint localization method was proposed for robust variable environments, but it has to rely on a Wi-Fi terminal with a transmitter and receiver module. Although we have analyzed AP's discrimination capability in [28], no localization algorithm was provided. Hence, the research of [28] was extended in this article, a new FDDC-WKNN technique is redefined by combining the policies of improving FD's discrimination capability to PD and WKNN algorithm, which can provide a solution for the aforementioned problem about APs' state.

The scenes considered is APs-rich indoor environments and APs' states are unknown. We conduct a variety of experiments to observe the effects of the various states of APs on localization accuracy, and find that the PD of the pairwise position indicated by RSSD depends on three factors. Then, based on the analysis of the relationship between the RSSD from an AP and a position pair, we provide the corresponding solution policies to improve FD's discrimination capability to PD from above three factors. At last, we integrated all these perspectives with WKNN in a unified solution and proposed an FDDC-WKNN algorithm.

The main work and contributions of the article can be summarized as follows.

- The root causes of limiting the discrimination capability of FD related to APs' state are investigated. We deeply analyzed and recognized three factors of APs affecting FD's discrimination capability, which is 1) the distance between AP and the sample position; 2) the AP's direction to the pairwise positions; 3) the health state of the APs.
- For the above three problems, we provided corresponding solution policies to improve FD's discrimination capability, which include a threshold to eliminate abnormal APs, a discrimination correction quantity, and a priority weight.
- Ultimately, by integrating the solution policies with WKNN, we advanced a redefined indoor localization technique, i.e. FDDC-WKNN, which has strong robustness and stability to the state of AP in APs-rich environments without knowing APs.

The rest of the manuscript is organized as follows. Section 2 introduces the framework of FDDC-WKNN and presents preliminary knowledge and some symbols definitions, and in section 3, we conduct some observation to find APs' state affecting FD's discrimination capability and provide corresponding solutions. FDDC-WKNN algorithm process is described in detail in Section 4, and in Section 5, extensive experiments and simulations are executed. At last, the conclusions of the article are presented in Section 6.

2. Indoor Fingerprint Localization Model and Preliminary

In this section, we first introduce the framework of the proposed FDDC-WKNN indoor fingerprint localization system, which is also applicable to other FDDC-based indoor localization systems. Then some preliminary knowledge and symbols definitions are provided to make the expression clearer later.

2.1. The Framework of the Indoor Fingerprint Localization Model

The localization process of a typical indoor fingerprint localization technique usually includes two phases, namely offline site survey and online fingerprint matching. During the offline phase, for each Reference Point (RP), the RSS values from different APs are collected and recorded as an RSS vector. Add the coordinates of the RP to the RSS vector to form the fingerprint of the RP. Finally, the fingerprints of all RPs form a fingerprint matrix and are stored in the fingerprint database. During the online phase, a user sends a RSS vector at a Test Position (TP) to the localization server, through the localization technique based on the fingerprints in the database, the server returns the user's estimated position.

In this article, we try to improve the WKNN algorithm based on FDs' discrimination capability enhancement to be a localization technique FDDC-WKNN at the server. The framework of the indoor fingerprint Localization system is shown in Figure 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

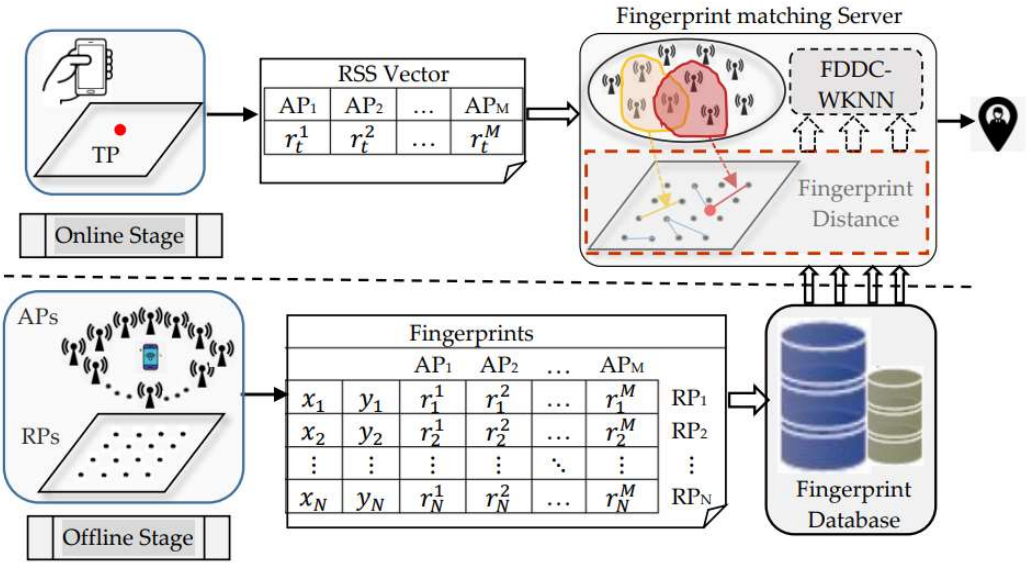


Figure 1. Framework of indoor fingerprint localization system.

2.2. Preliminary

For the convenience of subsequent expression, preliminary knowledge and some important symbols definitions used in the article are described and introduced in this section.

2.2.1. Log-normal Distance Path Loss Model

Theoretically, RSS decays logarithmically with propagation distance in terms of the propagation law of wireless signals. That is, when the transmitter sends a signal to the receiver, the path loss at the receiver ends logarithmically with the distance between the transmitter and the receiver. Here, to express signal propagation loss in a complex indoor scene, the Log-normal Distance Path Loss model is considered, see Formula (1), and then, the received signal strength at the receiver can be gained by Equation (2).

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

$$r(d) = P_R(d_0) - 10nlog\left(\frac{d}{d_0}\right) + X_\sigma, \quad (2)$$

Here, all distances are measured in meters and the parameters in (1) and (2) are explained as follows.

d : the distance from the transmitter to the receiver;

$r(d)$: the received signal strength at the receiver;

$P_L(d)$: the path loss at the receiver;

d_0: the reference distance, usually 1 m;

$P_L(d_0)$: the path loss at the reference distance;

n: the path loss exponent, i.e., the growth rate of path loss with distance which is based on the different environments and building types. mostly between 2 and 4.

X_σ: Gaussian noise, denotes the complex indoor environment effects.

$P_R(d_0)$: $P_T - P_L(d)$ and P_T is the transmitter's transmission power.

2.2.2. The Standard for Measuring the Relationship Between Two Vectors

Since utilizing FD of pairwise positions to discriminate the PD between them is a prerequisite for indoor fingerprint localization technologies, it is necessary to discuss the discrimination capability between two vectors, which means that we need to know under what conditions one vector can be

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

used to represent another vector well. Here, two criteria are considered, that is correlation and consistency. Next, let two vectors be denoted as X and Y, respectively.

- Correlation

In the field of natural sciences, the Pearson correlation coefficient between two variables is widely used to test the degree of correlation, with a value varying from -1 to 1. If the coefficient is between 0.8 and 1.0, the relationship between the two variables is a very strong correlation, and if the coefficient is between 0.6 and 0.8, the relationship between the two variables is a strong correlation. For X and Y, the Pearson correlation coefficient can be calculated by Formula (3). then the coefficient can be used to imply the correlation between FDs and PDs, and it is reasonable to be used to indicate the degree of linear correlation between them.

$$\lambda = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}, \tag{3}$$

where $\bar{X}$ and $\bar{Y}$ represent the mean values of X and Y, respectively.

- Consistency

Besides, it is reasonable that if the X_i is greater than X_j, the corresponding Y_i should also be greater than Y_j, which can be called the degree of a data sequence. Similarly, if the consistency is between 0.8 and 1.0, the relationship between the two variables is a very high consistency, and if the coefficient is between 0.6 and 0.8, the relationship between the two variables is a high correlation. Usually, the consistency coefficient of X and Y is defined as:

$$c = \frac{\sum_{1<i,j<n,i<j} sign[(X_i - X_j) * (Y_i - Y_j)]}{\sum_{1<i,j<n,i<j} 1}, \tag{4}$$

where

$$sign(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0' \end{cases} \tag{1}$$

From the above formula, there are some unreasonable situations for indoor fingerprint localization techniques. For example, if the PD of two positions is equal while the RSSD is very small, it is reasonable and practical to consider the data sequence between them to be consistent, but it is not consistent with the above formula. In this article, we use the modified formula of consistency coefficient in [28].

The aforementioned double coefficients are mainly used to test the relationships between FDs and PDs. In case the relationships have a very strong correlation and very high consistency, by adopting FDs to discriminate PDs in the indoor localization method, not only the uncertainty of signal propagation in the complex indoor environment can be greatly reduced, but also the localization accuracy can be improved and the localization error can be reduced.

2.3. Symbols Definitions

For one position i with coordinate (x_i, y_i), the RSS vector collected from m APs is represented by R_i:

$$R_i = [r_i^1, r_i^2, \dots, r_i^m], \tag{6}$$

where $r_i^k, (1 \leq k \leq m)$ is the RSS value that is received from the k-th AP. Combining the position coordinates and the RSS vector, we can get the fingerprint F_i of the position i:

$$F_i = [x_i, y_i, r_i^1, r_i^2, \dots, r_i^m], \tag{7}$$

For the pairwise positions (i, j) consisting of two different positions in the indoor localization area, whose RSS vectors are R_i and R_j respectively, and whose coordinates are (x_i, y_i) and (x_j, y_j) respectively, define the following variables.

- *RSSD* of the pairwise positions (i, j):

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

$$\Delta R_{ij} = \left[\Delta r_{ij}^1, \Delta r_{ij}^2, \ldots, \Delta r_{ij}^m\right], \tag{8}$$

where $\Delta r_{ij}^k = \left|r_i^k - r_j^k\right|$ indicates the RSSD of the k-th AP at positions i and j.

- *PD*:

$$d_{ij}^p = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}, \tag{9}$$

  Obviously, $d_{ij}^p = d_{ij}^p$.

- *FD* of the pairwise positions (i, j):

$$d_{ij}^f = \sqrt{\sum_{k=1}^{m} \left(\Delta r_{ij}^k\right)^2}, \tag{10}$$

Noted that, from Equation (10), it can be seen that the contribution of different RSSD to the discrimination capability of FDs is different. The bigger contribution of the RSSD Δr_{ij}^k means the more suitable AP is selected.

In order to measure the whole discrimination capability of the FD to the PD in the localization area, the whole FD vector and PD vector are defined. For the fingerprint of n RPs in the fingerprint database, write the PD vector of all pairwise positions as:

$$D^p = \left[d_{ij}^p\right], (0 < i \le n, j \le n, i < j), \tag{11}$$

Similarly, the FD vector of all pairwise positions is:

$$D^f = \left[d_{ij}^f\right], 0 < i \le n, j \le n, i < j, \tag{12}$$

In this article, the relationships between vectors D^p and D^f are measured in the correlation and consistency, which are used to measure the discrimination capacity of FD against PD.

3. Observation and Enhancement Policies for AP's Discrimination Capability

In the current highly developed social environment of information technology, our smart devices can easily receive high-quality signals from a dozen or more APs in public indoor environments.

These abundant APs are usually installed for the function of communication by different individuals, which are located at various positions and have various states including removed, impaired or unstable, etc. If there exist abnormal APs, it will inevitably lead to wrong FDs of pairwise positions, resulting in large localization errors. Therefore, utilizing such abundant existing APs to select the suitable APs to improve the discrimination capability of FD has become an important factor affecting localization accuracy.

From the definition of FD, we have to emphasize the discrimination capability of an AP's RSSD to fingerprint a specific position. So, next, according to an in-depth analysis of the discrimination capability of a single AP to a position pair, we aim to provide different AP RSS correction and selection ways, to greatly improve the discrimination capability of FD to a PD.

3.1. Observations for Discrimination Capability

Indoors, signal attenuation and RSS fluctuation are severe, which are caused by signal reflection and refraction, or multipath fading and indoor noise. Especially, they are closely related to the current environment. Hence, RSS is seriously affected by the indoor environment. In addition to the above factors, due to the influence of the distance, direction and states of the APs, it often occurred that the identical RSSD implies different PD between two positions.

In this article, the term discrimination capability is used to describe one AP's contribution to FD's capability to distinguish the PD of pairwise position.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

3.1.1. Diverse FDs' Discrimination Capability Caused by Different Distances APs

It can be seen from Formula (2), $RSS \propto -\log(d)$ which indicates that $\frac{\Delta r}{\Delta d} \propto \frac{1}{d'}$ where Δr is the RSSD and Δd denotes the corresponding distance difference. In other words, an identical Δr means a larger PD difference Δd at a farther position, or a smaller Δd at a closer position. Similarly, an identical Δd means a bigger RSSD Δr at a closer position, or a smaller Δr at a farther position. Figure 2 illustrates the RSS spatial distribution. It can be seen that the same RSSD of $\Delta r = 20$ corresponds to the different difference in PD of two pairwise positions, depending on the distances of the specific AP.

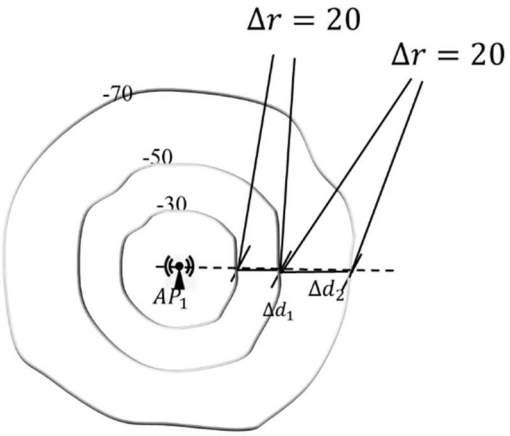


Figure 2. Discrimination capability diversity for APs with different distances.

In order to display the effect of the change of AP distance to discrimination capability, for the identical $\Delta d = 2m$ at a different distance from an AP, the correspondence Δr values are shown in Figure 3.

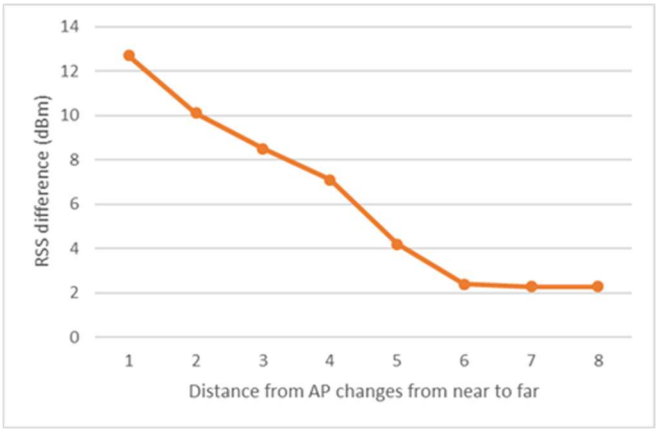


Figure 3. The different RSSDs for the same PD.

In a word, the PDs of pairwise positions indicated by RSSDs depend on the receivers' distance, leading to varying discrimination capability for various positions.

3.1.2. Diverse FDs' Discrimination Capability Caused by APs of Different Directions

Inherent constraints constrained by the law of radio signal propagation; APs have diverse discrimination capabilities to different pairwise positions. That is, an identical Δr can imply various PDs, see Figure 4., for any two positions, one is on the ring with RSS=-50dBm, and the other is on the ring with RSS=-70dBm, the value of RSSD Δr is always 20, but the PD between them varies greatly. According to simple geometry, it's easy to know that the PD of position pair (A, B) is the smallest

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

and (A, C) is the biggest. Meanwhile, the most accurate discrimination capability of RSSD $\Delta r = 20$ should be for position pair (A, B), which is related to the direction of the APs.

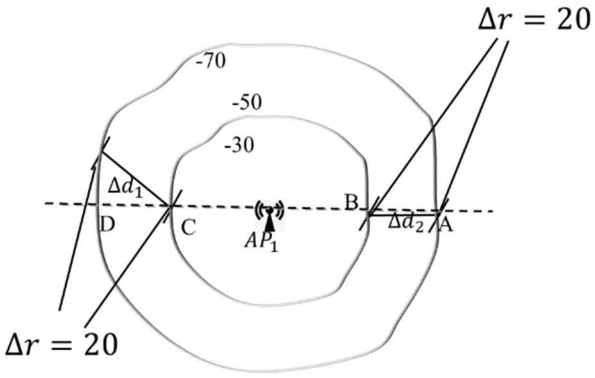


Figure 4. The discrimination capability diversity for APs with different directions.

In other words, see Figure 5, assume that the pairwise position (A, B) is fixed, i.e. $d_{AB} = d_0 \coloneqq \Delta d(d_0$ is a constant). Take position A as the center, and make a circle with a radius smaller than d_0.of d_{OA}. Suppose an AP is located at O, then d_{OA} is also a constant and $d_{OA} < d_0$, with O as the center, drawing two circles with a radius of d_{OA} and d_{OB}, respectively. Connecting OB and OA, which will intersect with the two circumferences at points A' and B', respectively, and θ $(0 \leq \theta \leq \pi)$is the included angle between OA and OB.

It is obvious that θ changes with the location of the AP, but the correspondence between the RSSD Δr_{AB} and radius difference $d_{OB} - d_{OA}$ remain unchanged, so set $d_{OB} - d_{OA} \coloneqq \Delta r$. Therefore, by studying the relationship between Δr and θ, we can get that when the orientation of AP changes concerning the position pair, the corresponding discriminate capability of RSSD changes rules. For $\Delta AA'B$, the following equation is gained:

$$\Delta r = d_0 - 2d_{OA} sin\frac{\theta}{2}, \tag{13}$$

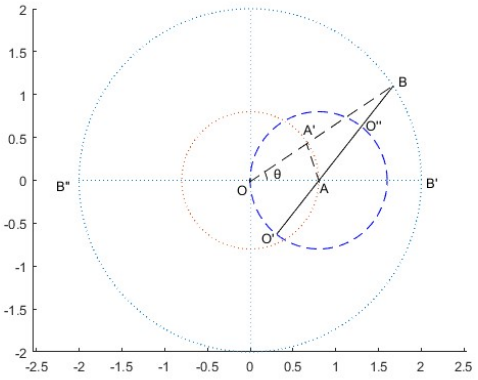


Figure 5. APs at different directions for a fixed position pair (A, B).

We can obtain that the same value Δd corresponds to different Δr, depending on the angle θ between line OA and OB, which means APs at different directions have different discrimination capabilities for the same position pair (A, B). As seen from Figure 5., when one AP is located at point O' and point O'', the discrimination capability of RSSD is optimal and poorest for position pair (A, B).

Remark that if d_{OA} is a very small value, which means that the space between AP and point A is tiny enough, Δd is about equal to the PD d_0. Then, the RSSD can characterize all PDs (A, B) well.

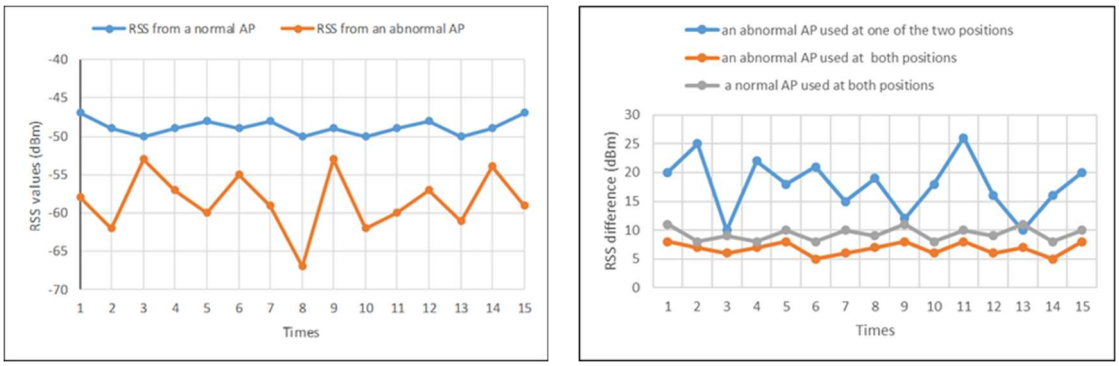
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

In a word, the PDs of the pairwise positions indicated by RSSDs depend on the receivers' direction, leading to varying discrimination capability for various directions.

3.1.3. Diverse FDs' Discrimination Capability Caused by APs of Different Health States

In order to make the localization more accurate, in addition to the above two aspects, the unstable APs should not be ignored. The unstable APs can be roughly divided into slight abnormality and complete failure (such as broken/removed/newly emerged). Slight abnormality means that the AP can transmit, but due to the device being impaired, the signal transmitted by the device is weak or unstable. Here, considering the complexity and variability of the indoor environment, for the obstruction of the new facilities or walls results in the weak signal of an AP at a specific location, we also treat the AP as slightly abnormal at the position. For complete failure/removed/newly emerged APs, it means that the device cannot transmit signals at all.

In order to observe the effect of normal APs and abnormal APs, we carry out an experimental analysis of the RSS values of a specific position and the RSSD of a position pair. First, for the sampling position that is fixed 3 meters away from AP, we sample RSS values from a normal AP and an abnormal AP respectively, results shown in Figure 6 (a). From the figure, it can be seen that the RSS values of the abnormal AP are obviously lower than the RSS value of the normal AP. In addition, the RSS values of the abnormal AP are extremely unstable compared to the values of normal AP. Second, for the sampling position pair with 2.5m PD, we calculate the RSSD of the position pair, see Figure 6 (b). As seen, when the AP sampling at one position is normal and at the other position is abnormal, the RSSD is obviously larger than that of normal AP and they are extremely unstable. When the AP sampling at both positions is abnormal, the RSSD is slightly smaller than that of normal AP and can be ignored when calculating FD.



(**a**) (**b**)

Figure 6. RSS value analysis. (a) The sampling point at 4m away from AP. (b) RSSDs for position pair with 2.5m.

In a word, the PDs of pairwise positions indicated by RSSDs are closely related to the AP's state, leading to diverse FD's discrimination capability for different healthy states.

3.2. FD's Discrimination Capability Enhancement Policies

In this section, different discriminatory policies are discussed based on the above discussion and analysis. we first discuss the strategy to identify the abnormal APs.

3.2.1. Abnormal APs Identification

Duo to abnormal APs is the main cause generating big errors in indoor fingerprint localization, we have designed two policies to reduce the impact of abnormal APs, one for slight abnormality APs and the other for complete failure/removed/newly emerged APs.

Since the abundance of APs in the indoor environment and the possibility of abnormal APs, it is unnecessary and also unrealistic to utilize the RSSDs of all APs equally. Whether for the fingerprint

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

in the fingerprint database or online collected the RSS of TP, set a threshold for RSS values to exclude APs with weak signal strength.

For position i, the RSS vector collected from m APs is $R_i = [r_i^1, r_i^2, \dots, r_i^m]$, the RSS value r_i^k is defined as:

$$r_i^k = \begin{cases} r_i^k & r_i^k \geq \mathcal{R} \\ 0 & r_i^k < \mathcal{R} \end{cases} \tag{14}$$

where $\mathcal{R}$ is a threshold, only the RSS values of APs with RSS values greater than $\mathcal{R}$ will be used to participate in the following steps, otherwise, RSS values will be set to 0. Through this method, slightly abnormal APs with small RSS values ($< \mathcal{R}$) in the environment can be eliminated. Furthermore, from Equation (8) and (10), we can find that the component Δr_{ij}^k of RSSD directly impacts the FD d_{ij}^f. An AP is normal at one position but is a complete failure or removed or newly emerged at the other position, which can generate a significant Δr_{ij}^k and affect the discrimination capability of FDs. To avoid this situation, make the following rule for Δr_{ij}^k:

$$\Delta r_{ij}^k = \begin{cases} \left| r_i^k - r_j^k \right| & r_i^k \neq 0 \; and \; r_j^k \neq 0 \\ 0 & r_i^k \neq 0 \; or \; r_j^k \neq 0 \end{cases}, \tag{15}$$

We assume that there always exist rich APs in indoor scenes, then through the above policies, abnormal APs, especially the complete failure or removed or newly emerged APs in the environment can be eliminated almost.

3.2.2. Discrimination Correction Quantity

Here, we define a impact factor to quantitatively differentiate the RSSD of each AP for a specific position pair. Since a position pair includes two specific positions, the discrimination factor relates to the impact of AP at each position, it is necessary first to quantify the impact of each AP at each location which we term as an impact factor, and then to define correction quantity based on impact factor.

According to the LDPL model (2) and the estimation of PDs between AP and a position, for the k-th AP to the i-th position, the impact factor is calculated as follows:

$$\gamma_i^k = \frac{1}{\hat{d}_i^k} = 10^{\frac{r_i^k - P_R(d_0)}{10n}} \tag{16}$$

where $\hat{d}_i^k$ is the estimated distance between the k-th AP and the i-th position.

Considering that a certain AP has a large RSS value due to RSS fluctuation, which makes the other representative APs negligible, we introduce the way in literature [28] to redefine the impact factor. For coding with noisy RSSs, the impact factor is adjusted as follows:

$$\gamma_i^k = \begin{cases} 10^{\frac{r_i^k - P_R(d_0)}{10n}} & if \; r_i^k \leq r_0 \\ \frac{1}{a}\left(1 + e^{-2\left(\frac{r_i^k + 100}{10} - c\right)}\right)^{-1} & otherwise \end{cases}, \tag{17}$$

where r_0 is a flexible value, it can be given by empirical, e.g., -50dBm, a and c are constant parameters, which can be determined based on n such that γ_i^k is continuous at r_0.

From formula (17), because the reciprocal of PD is consistent with the monotonicity of the LDPL model, the basic rule is that the closer APs are, the stronger RSSs are. Due to the RSSD is the difference of RSS values between two positions, considering the uncontrollability of AP states in indoor environments, we believe that the larger the RSS, the greater the interference for RSSD. Therefore, the correction quantity is defined as follows.

$$\varphi_i^k = \Phi \cdot \gamma_i^k \tag{18}$$

where Φ is a constant between 3dBm-6dBm, according to the environment conditions, approximately equal to the value of signal fluctuation at a position.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

3.2.3. Priority Weight

From Formula (13) in section 3.1.2, we can gain a rule: for APs in different directions, the smaller the angle θ, the larger the RSSD Δr, and the stronger the discrimination capability of FD, which means that the APs who get the larger RSSD has greater contribution to FD's discrimination capability. Therefore, we grant each RSSD of each pairwise position a priority weight τ_{ij}^{k} as follows, which determines the priority and contribution of its participation in FD, see Equation (19). As seen, the larger RSSD corresponds to a higher priority weight.

$$\lambda_{ij}^{k} = \frac{1}{1 + e^{-\Delta r_{ij}^{k}}}, \tag{19}$$

It is clear that λ_{ij}^{k} is the S function of Δr_{ij}^{k}, so adding λ_{ij}^{k} to RSSD will not have a significant impact on the value of FD, avoiding affecting localization accuracy.

Overall, through the above provided discrimination capability improved policies, the APs with large contributions to FD's discrimination capability will be involved in the calculation of FD, and the more suitable the APs, the higher the degree of participation. Hence, FD with strong discrimination capability to PD is obtained, providing a guaranteed prerequisite for relevant fingerprint matching algorithms.

4. WKNN and FDDC-WKNN Algorithm

4.1. The Idea of KNN and WKNN Algorithms

The K-nearest neighbor (KNN) algorithm is one of the most widely used fingerprint matching algorithms. The idea of the KNN algorithm is to select the positions of K fingerprints closest to the current positional fingerprint to estimate the current position, which is simple, intuitive, and effective. KNN is one of the algorithms that use FD of pairwise positions to discriminate PD between them.

We assume that there are n RPs in an indoor area, and n fingerprints were collected in the offline phase, which is denoted as $[F_1, F_2, \cdots, F_n]^T$, where $F_i = [x_i, y_i, R_i]$, see Equation (6) and (7). In the online phase, the RSS vector of a TP is R_t. First, calculate all FDs $d_{ti}^{f}(0 < i \leq n)$ from position S to RPs; second, select *K* nearest RPs based on all FDs; finally, use the following Formula (20) to get the estimation value of S coordinates:

$$S =: (\hat{x}, \hat{y}) = \frac{1}{K}\sum_{i=1}^{K}(x_i, y_i) \tag{20}$$

The above formula indicates that the KNN algorithm sets the same weight to all K nearest neighbors. However, according to the propagation characteristics of indoor signals, it is known that the farther the distance, the stronger the signal fluctuation. For this reason, the WKNN algorithm sets different weights based on FD to emphasize that close neighbors contribute ability more, the estimation value of S coordinates can be expressed as follows.

$$S =: (\hat{x}, \hat{y}) = \frac{1}{\sum_{i=1}^{K} w_i}\sum_{i=1}^{K} w_i * (x_i, y_i) \tag{21}$$

where w_i is the weight related to distance. In this article we adopt the WKNN algorithm to be the indoor fingerprint algorithm, and the weights are given in the following formula.

$$w_i = \frac{1/d_{ti}^{f}}{\sum_{i=1}^{K} 1/d_{ti}^{f}} \tag{22}$$

Note that: Different K values mean that the number of nearest RPs selected is different, which can affect the accuracy of localization. The optimal K value often gained from experience or practice

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

4.2. The FDDC- WKNN Algorithms

Taking into account all the strategies and WKNN idea discussed above, a uniform WKNN solution based on improving FD's discrimination capability will be presented in this section, that is FDDC-WKNN. The presence of a large amount of abundant AP in the indoor environment is a prerequisite for this solution.

For an indoor area, given that there are N RPs and M APs and fingerprint databases have been built on the Server. In the online phase, when we have the RSS vectors R_t at a TP, send them to the server, and want to get the TP location $(\hat{x}, \hat{y})$. Based on the policies aforementioned, the algorithm steps on the server side are given as follows.

Step1. Set a threshold $\mathcal{R}$, modify the RSS value in R_t and the fingerprint data in dataset by Formula (14), and combine r_t and the RSS matrix of the fingerprint to a new matrix F'as follows, where the first row of F' is RSS vector R_t at TP, and the other N rows are RSS matrix at RPs.

$$F' = \begin{bmatrix} r_t^1 & r_t^2 & \cdots & r_t^M \\ r_1^1 & r_1^2 & \cdots & r_1^M \\ r_2^1 & r_2^2 & \cdots & r_2^M \\ \vdots & \vdots & \ddots & \vdots \\ r_N^1 & r_N^2 & \cdots & r_N^M \end{bmatrix} := [R_t \quad R_1 \quad R_2 \quad \cdots \quad R_N]^T, \tag{23}$$

Step2. Calculate discrimination factor γ_i^k for each RSS value and gain discrimination correction quantity matrix Γ as follows.

$$\Gamma = [\Gamma_t \quad \Gamma_1 \quad \Gamma_2 \quad \cdots \quad \Gamma_N]^T, \tag{24}$$

where

$$\begin{gathered} \Gamma_i = [\varphi_i^1 \quad \varphi_i^2 \quad \cdots \quad \varphi_i^M]^T, (i = t \; or \; 1 \le i \le N), \\ \varphi_i^k = \Phi \cdot \gamma_i^k \end{gathered} \tag{25}$$

Then, we gain the correction quantity by Γ times Φ and further calculate modified RSS data by F' minus discrimination correction quantity matrix $\Gamma * \Phi$, see Equation (26). Here, $\tilde{R}_1$ can be seen as the RSS value corrected by the discrimination factor.

$$F' - \Gamma = [R_t - \Gamma_t \quad R_1 - \Gamma_1 \quad R_2 - \Gamma_2 \quad \cdots \quad R_N - \Gamma_N]^T := [\tilde{R}_t \quad \tilde{R}_1 \quad \tilde{R}_2 \quad \cdots \quad \tilde{R}_N]^T := \tilde{F}, \tag{26}$$

Step 3. Calculate the RSSD of the TP to each RP, that is subtract each remaining row of the matrix $\tilde{F}$ from the first row to obtain each RSSD vector $\Delta R_{ti} (1 \le i \le N)$, then, an RSSD matrix is gained as follows.

$$\Delta R = [\Delta R_{t1} \quad \Delta R_{t2} \quad \cdots \quad \Delta R_{tN}]^T, \tag{27}$$

where

$$\Delta R_{ti} = (|\varphi_t^1 r_t^1 - \varphi_i^1 r_i^1|, |\varphi_t^2 r_t^2 - \varphi_i^2 r_i^2|, \cdots, |\varphi_t^N r_t^N - \varphi_i^N r_i^N|)(1 \le i \le N), \tag{28}$$

Step4. Use priority weight Formula (19) to calculate all priority weight of $\Delta R_{ti} (1 \le i \le N)$, we can get the following priority weight matrix Λ.

$$\Lambda = [\Lambda_{t1} \quad \Lambda_{t2} \quad \cdots \quad \Lambda_{tN}]^T, \tag{29}$$

where $\Lambda_{ti} = [\lambda_{ti}^1 \quad \lambda_{ti}^2 \quad \cdots \quad \lambda_{ti}^M]^T, (1 \le i \le N)$

Then, calculate the Hadamard product of Λ and ΔR, that is multiply the corresponding elements of Λ and ΔR as Equation (30). Then, we get the weighted RSSD $\Delta \tilde{R}$ by priority weight.

$$\Lambda \odot \Delta R = [\Lambda_{t1} \Delta R_{t1} \quad \Lambda_{t2} \Delta R_{t2} \quad \cdots \quad \Lambda_{tN} \Delta R_{tN}]^T := [\Delta \tilde{R}_{t1} \quad \Delta \tilde{R}_{t2} \quad \cdots \quad \Delta \tilde{R}_{tN}]^T := \Delta \tilde{R}, \tag{30}$$

where

$$\Delta \tilde{R}_{ti} = (\lambda_{ti}^1 |\varphi_t^1 r_t^1 - \varphi_i^1 r_i^1|, \lambda_{ti}^2 |\varphi_t^2 r_t^2 - \varphi_i^2 r_i^2|, \cdots, \lambda_{ti}^M |\varphi_t^N r_t^N - \varphi_i^N r_i^N|)(1 \le i \le N), \tag{31}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

Step 5. Calculate the FD of TP to each AP by Formula (10) to obtain an FD sequence D_t^f and expressed by following Equation (32).

$$D_t^f = \left(d_{t1}^f \quad d_{t2}^f \quad \cdots \quad d_{tN}^f\right), \tag{32}$$

where

$$d_{ti}^f = \sqrt{\sum_{k=1}^{M} \left(\lambda_{ti}^k \left|\varphi_t^k r_t^k - \varphi_i^k r_i^k\right|\right)^2}, \tag{33}$$

Step6. Select the smallest K elements in sequence D_t^f and calculate their weights according to Formula (22). Furthermore, according to Formula (21) to get the position of TP.

In addition, in order to obtain the correlation and consistency between PD and FD in subsequent experiments, we need to calculate the PD of TP to each AP by Formula (9) to obtain an PD sequence D_t^p and expressed as follows.

$$D_t^p = \left(d_{t1}^p \quad d_{t2}^p \quad \cdots \quad d_{tN}^p\right), \tag{34}$$

where

$$d_{ti}^p = \sqrt{(x_t - x_i)^2 + (y_t - y_i)^2} \tag{35}$$

5. Simulation and Experiments

In this section, we will evaluate the policies and algorithm proposed in this article by conducting some experiments in two scenes, that is simulation environments and real environments.

In particular, our technology is suitable for having a large number of APs in both scenes and there are phenomena such as APs being damaged, malfunctioning, or removed, etc.

5.1. Scenes Setup and Some Parameters Presets

Firstly, the setups and some presets of simulation and real scenes are introduced, respectively.

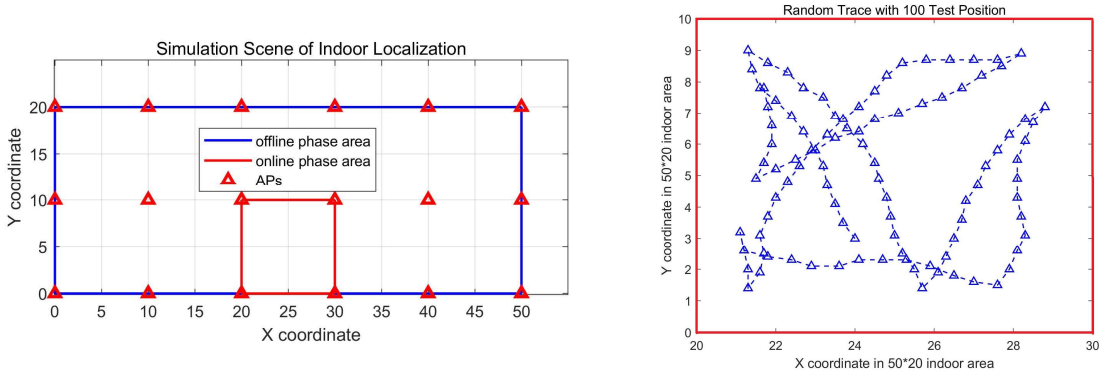
5.1.1. Introduce for Experimental Scene

- Scene 1: simulation scene

To simulate large-scale indoor space with rich APs, we simulated a 50 * 20 sized indoor area with 18 APs, and these APs are placed on grid points every 10 meters in the simulation area, which is depicted in Figure 7 (a). In our simulation process, RSS data are generated by ray tracing technology based on the LDPL model while considering wall reflection and noise indoors. Initially, we generated a simulation fingerprint dataset at position with an interval of 0.1m. Then we can get any offline fingerprint database from the dataset as needed. For example, an offline fingerprint database with an RPs interval of 1 m only needs to take values every 10 position points in the dataset, that is, 10 * 0.1=1m, and so on.

In our following simulation experiment, the offline fingerprint database is from the above fingerprints dataset by setting the RPs interval of the offline database to 1 m, that is there are 51*21 = 1071 RPs in the simulation area.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1



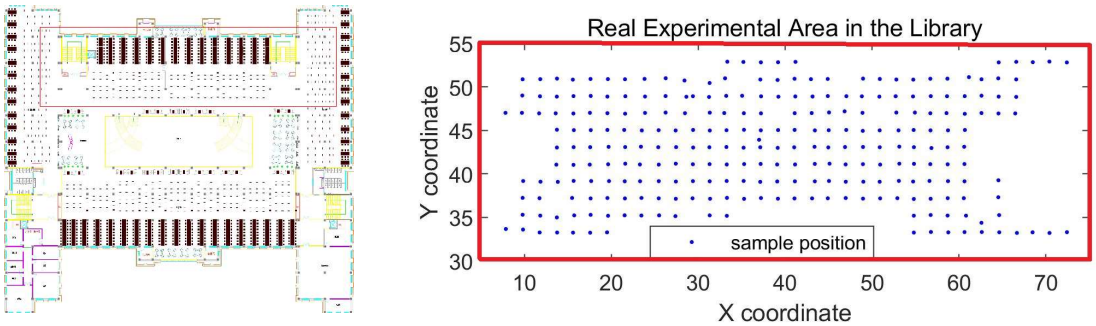
(a) (b)

Figure 7. Simulation scene of indoor localization. (a) Details of simulation area; (b) Random trace with 100 TPs.

In the online phase, we simulated a target moving in a 100 square meter room in the simulation area (see the red rectangle in Figure 7 (a)) and obtained a trace with 100 trajectory points, as well as RSS on each trajectory point, see Figure 7 (b). Meanwhile, we simulated two abnormal APs, one of which was removed or broken, and the other had a minor malfunction These 100 trajectory points are used as TPs for the indoor localization algorithm. In order to simulate people's living habits, the movement trajectory moves in space within 1 meter of the boundary.

- Scene 2: real library Scene

The real experimental scene was on the second floor of the school library in Dalian University, and the plan of which is shown in Figure 8 (a). The central position of this floor is the vacant part on the first floor, surrounded by an open bookshelf reading area, including a large number of open bookshelves and reading tables and chairs. The experiment area is located on the north side of the floor with about 1200m^2, in the red rectangle area of Figure 8 (a), and Figure 8 (b) is the enlarged experiment area, as can be seen, there are 261 RPs at an interval of 2m. In addition, there are a total of 83 APs that can work in the experimental area.



(a) (b)

Figure 8. Real Scene of Indoor Localization. (a) The plan of the second floor in the school library. (b) Real experiment area with APs.

In the offline phase, we used HUAWEI P9 smartphone to sample RSS from 83 APs to build a fingerprint database. In the online phase, to reduce the workload of data sampling, we randomly selected 50 points from these 261 to be TPs and set 8 abnormal APs with 4 APs removed or breakdowns and 4 APs with slight abnormality.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

Overall, in order to have a clearer understanding of the two scenes, the setups for both scenes are shown in Table 1.

Table 1. Setups for Two Scene.

Scenes	Sample Dimensions	Experiment Dimensions	Number of APs	APs' Positions	Abnormal APs	Number of TPs
Scene 1	50m*50m	10m*10m	18	Grid Points /10m	2	100
Scene 2	1200m^2	1200m^2	83	unknown	8	50

5.1.2. Parameters Presetting

In the proposed algorithm, many parameters need to be set to experiment. Here, considering the environments of two scenes, the parameters are determined based on depending on experience or specific conditions respectively, as listed in Table 2.

Table 2. Presets for Parameters in the experiment.

Parameters	Scene 1	Scene 2
Path loss exponent n	2	/
Reference distance d_0	1	/
Power at d_0	-39	/
Parameters r_0, a, cin impact factor	-55,3.4,4.288	-55,3.4,4.288
Threshold $\mathcal{R}$	-70	-90
Fluctuating quantity Φ	3	5

Due to the RSS being sampled directly in the real scene, the first three parameters are not required to be set.

5.2. Simulation and Analysis

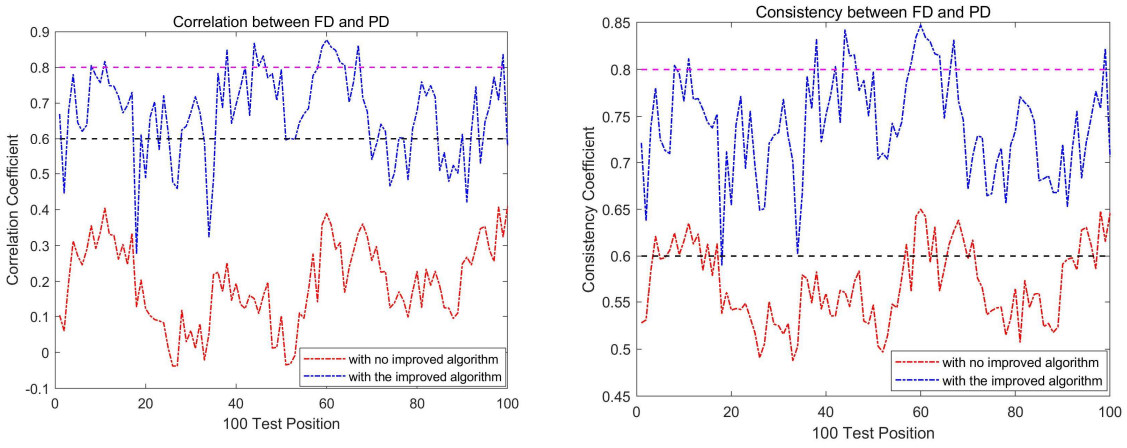
Aiming to measure whether the proposed policies can improve the whole discrimination capability of FD to PD, we conducted some experiments to verify the correlation and consistency changes between FD and PD before and after using FDDC-WKNN. For the simulation environment, we first verified the effect of proposed policies on the discrimination capability of FD to PD. Then, for a random trace with 100 positions, we find the optimal value of K, and with the optimal K, we simulate the localization positions and derive the CDF curves to demonstrate the localization performance.

For the real scene in the Dalian University library, we also find the optimal value of K for 50 random TPs, and with the optimal K, we gain the localization results similarly.

5.2.1. Discrimination Capability of FD to PD

By calculating PD sequence D_t^p by Equation (35) and FD sequence D_t^f by Equation (35) between 100 TPs and 81 RPs in the localization area, we gain the correlation coefficient and consistency coefficient between each TP and RPs. From the results shown in Figure 9, we can see that both the correlation coefficient and the consistency coefficient are greatly improved. After implementing the policies proposed in this article, strong correlation increased from 0 to 76%, and consistency increased from 26% to 99%, which indicates the discrimination capability of FD to PD has been greatly improved.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1



(**a**) (**b**)

Figure 9. Correlation and Consistency between PD and FD. (a) Correlation coefficient before and after using the proposed policies. (b) Consistency coefficient before and after using the proposed policies.

To further examine the impact of abnormal AP on FDs' discrimination Capability, we repeated the experiment of calculating D_t^p and D_t^f under diverse abnormal AP conditions to obtain the correlation coefficient and consistency coefficient between PD and FD. For clarity, the experimental conditions are recorded as following situations, and the results before and after implementing FDDC_WKNN are displayed in Table 2.

I: All APs are in a normal state

II: There is one slight abnormality AP

III: There is one complete failure AP

Table 2. Correlation and consistency between PD and FD before and after implementing the proposed policies.

	I		**II**		**III**		**II&III**	
Strong Correlation	16%	91%	16%	89%	0	78%	0	76%
High Consistency	72%	99%	78%	99%	27%	99%	26%	99%

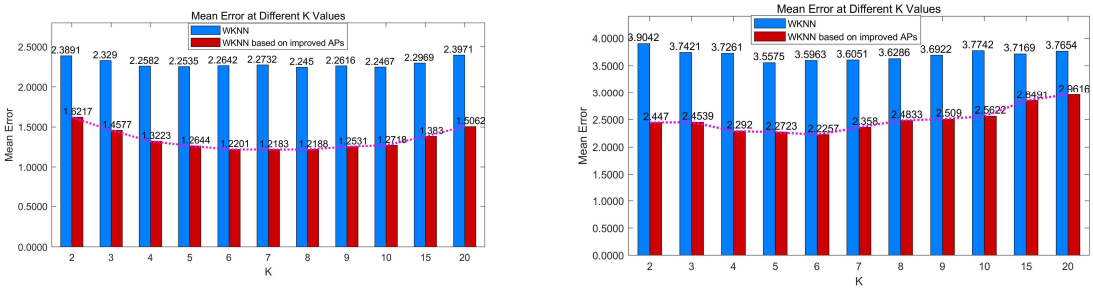
From the results in Table 2, it can be seen that the presence of completely failed APs in the localization area has the greatest impact on the correlation and consistency between FD and PD. Moreover, regardless of the conditions, the policies proposed in this article have improved the correlation and consistency between FD and PD to varying degrees.

5.2.2. Setting of K

In the same conditions of localization area, different K values in FDDC-WKNN can lead to different localization accuracy. We assumed that the K value with the lowest average errors is optimal. The optimal K value is usually not fixed and varies depending on the environment. We compared the average Errors under different K values for both simulation and real scenes to obtain the corresponding optimal K value.

As shown in Figure 10, the optimal K value is 7 in the simulation scene and 6 in the real scene. However, we found that the average localization errors did not change significantly with different K values, as our policies have greatly improved the FD's capability to discriminate PD, thereby reducing the impact of K value changes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

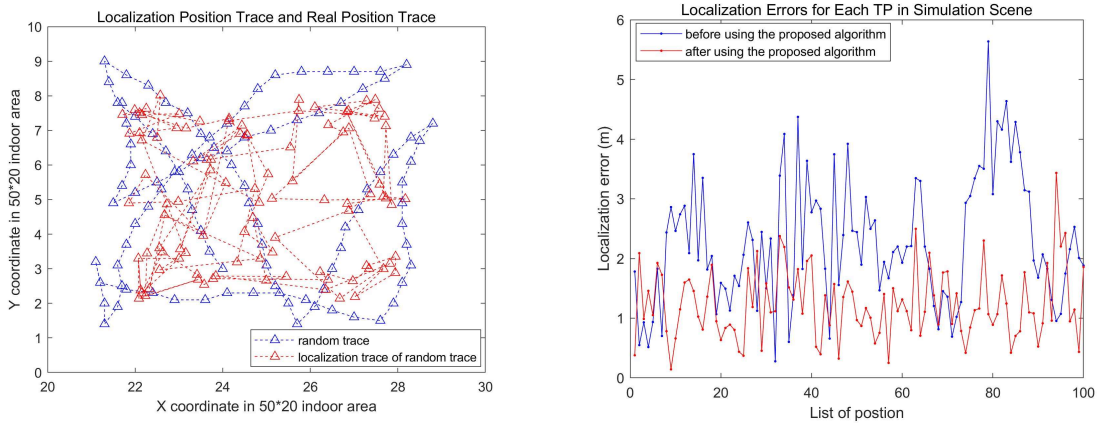


(**a**) (**b**)

Figure 10. Average localization errors under different K. (a) In simulation scene. (b) In the real library scene.

5.2.3. Localization Examination by FDDC-WKNN under the Optimal K

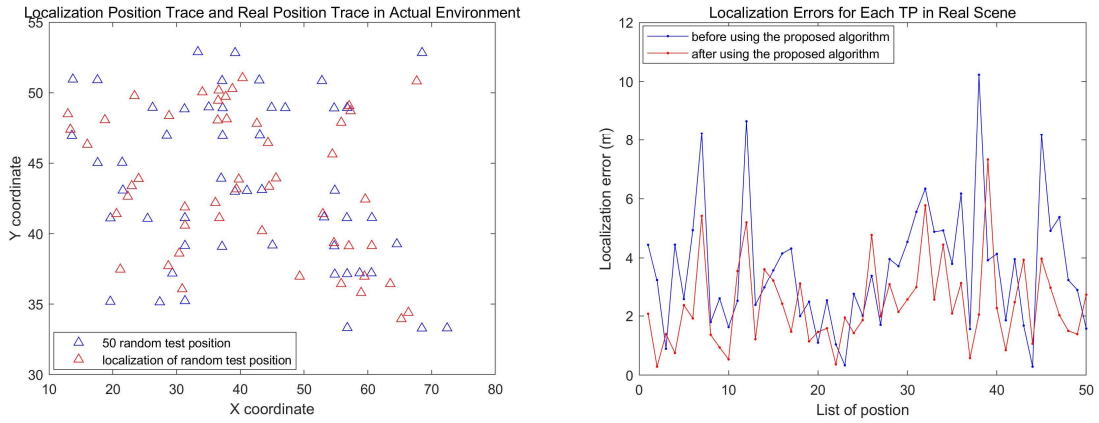
For $K = 7$ in the simulation scene with 100 TPs and $K = 6$ in the real library scene with 50 TPs, we apply the FDDC-WKNN algorithm, and the localization results are demonstrated in Figure 11 and Figure 12, respectively.



(**a**) (**b**)

Figure 11. Localization results applying FDDC-WKNN for simulation scene with 100 TPs. (a) The original position and localization results of TPs. (b) Localization errors of each TP.

In the simulation scene, before using FDDC-WKNN algorithm, due to the influence of abnormal APs, the localization errors were large and fluctuated greatly. Based on the FDDC-WKNN algorithm, it can be seen from Figure 11 (b) that the localization errors have significantly decreased and are basically stable.

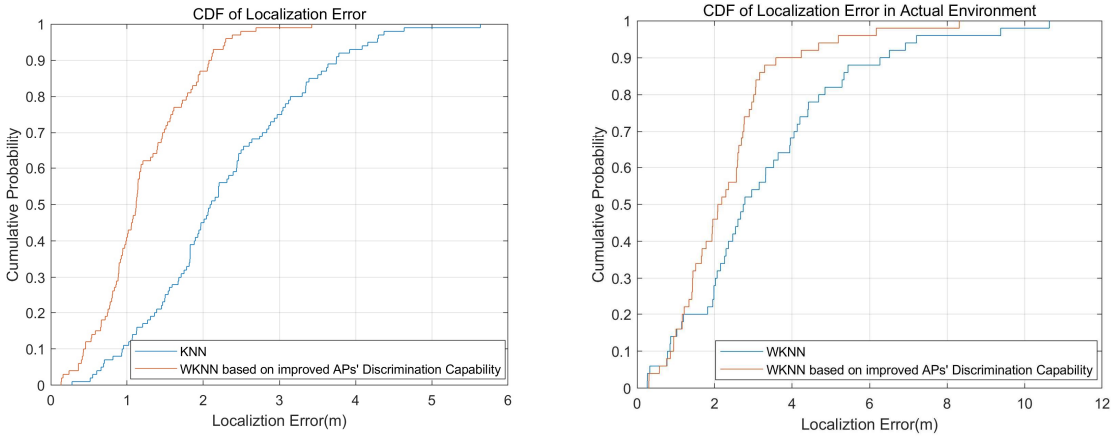


(**a**) (**b**)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202401.1075.v1

Figure 12. Localization results applying FDDC-WKNN for the real library scene with 50 TPs. (a) The original position and localization results of TPs. (b) Localization errors of each TP.

In the real library scene, the localization results are basically consistent with the simulation scene, see Figure 12, but due to the complexity and variability of the real environment, the localization errors are larger and the error fluctuation is also greater. Moreover, Figure 13 presents a comparison of the CDF of WKNN and FDDC-KNN, it is evident that FDDC-WKNN outperforms WKNN.



(**a**) (**b**)

Figure 13. CDF of localization errors for WKNN and FDDC-WKNN. (a) In the simulation scene. (b) In the real library scene.

In summary, we have examined the WKNN algorithm based on the improving FDs' discrimination capability proposed in this article through multi-directional experiments. All experimental results indicate that the proposed approach can improve FD's capability to discriminate PD in environments with abnormal APs, thereby reducing localization errors and improving localization accuracy. In addition, the proposed algorithm performs stably in the presence of interference in the environment, proving its robustness.

6. Conclusions

In this article, by defining a threshold, a correct quantity, and a priority weight, the capability of FD to discriminate PD is greatly improved for complexity and AP-rich indoor environments. Furthermore, combined with WKNN, a new indoor fingerprint localization technique referred to as FDDC-WKNN was provided. FDDC-WKNN algorithm is particularly suitable for indoor scenarios with uncontrollable APs and has a stable localization performance.

The discrimination capability of FD to PD is the key foundation for most indoor localization methods based on FD. Next, the algorithm will be implemented, verified, and improved in the indoor localization process. Also, improving the accuracy of localization with the other various intelligent algorithms and auxiliary facilities, studying different strategies based on various environment requirements and so on are the future study directions.

Author Contributions: Conceptualization, B.W.; methodology, J.L., Z.W. and Q.Y; investigation, J.L.; resources, Z.W.; data curation, Q.Y.; writing—original draft preparation, B.W. and Q.L.; writing—review and editing, B.W. and J.L.; visualization, B.W. and Q.L.; supervision, Z.W. and Q.Y.; project administration, J.L., R.L. and B.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Hebei Province Higher Education Science and Technology Research Key Project, grant number ZD2022076; Scientific Research Project of Tianjin Municipal Education Commission, grant number 2022ZD018; and National Natural Science Foundation of China, grant number 61702071.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable

Acknowledgments: Authors express their gratitude to all reviewers to improve the work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. J. H. Seong; D. H. Seo, Selective Unsupervised Learning-Based Wi-Fi Fingerprint System Using Autoencoder and GAN, *IEEE Internet of Things Journal* **2020**, Vol. 7, No. 3, pp. 1898-1909.
2. P.Roy; C.Chowdhury, A survey on ubiquitous WiFi-based indoor localization system for smartphone users from implementation perspectives, *CCF Trans. Pervasive Comp. Interact* **2022**, 4, 298–318.
3. A. Famili; A. Stavrou; H. Wang and J. -M. J. Park, RAIL: Robust Acoustic Indoor Localization for Drones, 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring), Helsinki, Finland, 19-22 June 2022.
4. L. Santoro; M. Nardello; D. Brunelli and D. Fontanelli, UWB-Based Indoor Positioning System With Infinite Scalability, *IEEE Transactions on Instrumentation and Measurement* **2023**, vol. 72, pp. 1-11.
5. M. Merenda; L. Catarinucci; R. Colella; D. Iero; F. G. D. Corte and R. Carotenuto, RFID-Based Indoor Positioning Using Edge Machine Learning, *IEEE Journal of Radio Frequency Identification* **2022**, vol. 6, pp. 573-582.
6. Rathnayake, R.M.M.R.; Maduranga, M.W.P.; Tilwari, V.; Dissanayake, M.B, RSSI and Machine Learning-Based Indoor Localization Systems for Smart Cities *Eng* **2023**, 4, 1468-1494.
7. Y. Xiao; S. Zhang; J. Cao; H. Wang and J. Wang, Exploiting distribution of channel state information for accurate wireless indoor localization, *Computer Communications* **2017**, vol. 114, pp.73-83.
8. Liu, F.; Liu, J.; Yin, Y.; Wang, W.; Hu, D.; Chen, P.; Niu, Q., Survey on wifi-based indoor positioning techniques. *IET communications* **2020**, 14(9), 1372–1383.
9. Obeidat, H.; Shuaieb, W.; Obeidat, O.; Abd-Alhameed, R., A review of indoor localization techniques and wireless technologies. *Wireless Personal Communications* **2021,**119, 289–327.
10. Biswas, D.; Barai, S.; Sau, B., New RSSI-fingerprinting-based smartphone localization system for indoor environments. *Wireless Netw* **2023**, 29, 1281–1297.
11. C. Zhou; B. Wang; Y. Mo and Z. Zeng, MOCLoc: Emerging Online Collaborative Localization Enhanced by Multidimensional Scaling, *IEEE Transactions on Emerging Topics in Computational Intelligence* **2022**, vol. 6, no. 4, pp. 751-761.
12. Y. L. Xiao; S. G. Zhang; J. X. Wang, An indoor localization algorithm based on multidimensional scaling and region refinement, *Chinese Journal of Computers* **2017**, vol. 8, pp. 196-210.
13. Molla, J.P.; Dhabliya, D.; Jondhale, S.R.; Arumugam, S.S.; Rajawat, A.S.; Goyal, S.B.; Raboaca, M.S.; Mihaltan, T.C.; Verma, C.; Suciu, G, Energy Efficient Received Signal Strength-Based Target Localization and Tracking Using Support Vector Regression. *Energies* **2023**, 16, 555.
14. C. Li; Z. Qiu and C. Liu, An improved weighted k-nearest neighbor algorithm for indoor positioning, *Wireless Personal Communications* **2017**, vol. 96, PP. 2239-2251.
15. Lu, H.; Zhang, L.; Chen, H.; Zhang, S.; Wang, S.; Peng, H.; Zou, J. An Indoor Fingerprint Positioning Algorithm Based on WKNN and Improved XGBoost. *Sensors* 2023, *23*, 3952.
16. Q. G. Wen; Y. C. Liang; C. G. wu; A. Tavares and X. S. Han, Indoor localization algorithm based on artificial neural network and radio-frequency identification reference tags, *Advances in Mechanical Engineering* **2018**, vol. 10, no. 12, pp. 1-12.
17. Y. Lin; K. Yu; L. Hao; J. Wang and J. Bu, An Indoor Wi-Fi Localization Algorithm Using Ranging Model Constructed With Transformed RSSI and BP Neural Network, *IEEE Transactions on Communications* **2022**, vol. 70, no. 3, pp. 2163-2177.
18. Wisanmongkol, J., et al., An ensemble approach to deep-learning-based wireless indoor localization. IET Wirel. *Sens. Syst* **2022**, 12(2), 33–55.
19. Shang, S.; Wang, L., Overview of WiFi fingerprinting-based indoor positioning. *IET Commun*. **2022**, 16, 725–733.
20. D. Sikeridis; B. P. Rimal; I. Papapanagiotou and M. Devetsikiotis, Unsupervised crowd-assisted learning enabling location-aware facilities, *IEEE Internet of Things Journal* **2018**, vol. 5, no. 6, pp. 4699-4713.
21. Q. Li; W. Li; W. Sun; J. Li and Z. Liu, Fingerprint and assistant nodes based Wi-Fi localization in complex indoor environment, *IEEE Access* **2016**, vol. 4, pp.2993-3004.
22. T. Lan; X. Wang; Z. Chen; J. Zhu and S. Zhang, Fingerprint Augment Based on Super-Resolution for WiFi Fingerprint Based Indoor Localization, *IEEE Sensors Journal*, **2022**, vol. 22, no. 12, pp. 12152-12162.
23. H. Rizk; H. Yamaguchi; M. Youssef and T. Higashino, Laser Range Scanners for Enabling Zero-overhead WiFi-based Indoor Localization System, ACM Trans. *Spatial Algorithms Syst* **2023**, 9, 1, Article 4.

24. F. Zafari; A. Gkelias and K. K. Leung, A survey of indoor localization systems and technologies, *IEEE Communications Surveys & Tutorials* **2019**, vol. 2, no. 3, pp. 2568-2599.
25. M.Nabati; S.A.Ghorashi, A real-time fingerprint-based indoor positioning using deep learning and preceding states, *Expert Systems with Applications* **2023**,Volume 213, Part A, 2023,118889.
26. J. Hu and C. Hu, A WiFi Indoor Location Tracking Algorithm Based on Improved Weighted K Nearest Neighbors and Kalman Filter, *IEEE Access* **2023**, vol. 11, pp. 32907-32918.
27. Xiang, L.; Xu, Y.; Cui, J.; Liu, Y.; Wang, R.; Li, G. GM(1,1)-Based Weighted K-Nearest Neighbor Algorithm for Indoor Localization. *Remote Sens.* 2023, *15*, 3706.
28. B. Wang; Q. Yu; Z. Wang; J. Liu, Analysis of AP's Discrimination Capability for Fingerprint-based Indoor Localization Technology. 2023 IEEE International Conference on Smart Internet of Things(SmartIoT), Xining, Qinghai, China, August 25-27, 2023.
29. C.S.Wu; Z.Yang; Z.M. Zhou; Y. H. Liu and M. Y. Liu, Mitigating large errors in WiFi-based indoor localization for smartphones, *IEEE Transactions on Vehicular Technology* **2016**, vol. 66, no. 7.
30. J. Luo; X. Yin; Y. Zheng and C. Wang, Secure indoor localization based on extracting trusted fingerprint, *Sensors* **2018**, vol. 18, no. 2, 469.