



Research Article

Reliable Wi-Fi Indoor Localization in Case of AP Loss by Using Integrated Model Based on Signal Anomaly Detector and Signal Distance Corrector

Zheng Yao ^{1,2,3}, Huaiyu Wu ^{1,2,3}, Yang Chen,^{1,2,3} Zhihuan Chen,^{1,2,3}
and Xiujuan Zheng^{1,2,3}

¹Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan, China

²Institute of Robotics and Intelligent Systems, Wuhan University of Science and Technology, Wuhan, China

³School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan, China

Correspondence should be addressed to Huaiyu Wu; wuhy@wust.edu.cn

Received 7 January 2021; Revised 22 February 2021; Accepted 29 March 2021; Published 14 July 2021

Academic Editor: Shiping Wen

Copyright © 2021 Zheng Yao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

When developing a Wi-Fi indoor positioning system in a real-world environment, the problems we have to face are that some access points' signal strength fluctuates extensively or even loses contact due to the cybersecurity threats, leading to the fact that the indoor location system cannot get reliable application in a real-world environment. To solve this problem, we propose a new integrated model based on signal anomaly detector and signal distance corrector to provide reliable position estimation when the access points' signal is lost under cybersecurity threats. The signal anomaly detector improves recognition capability of the uncertain signal and noise, while the signal distance corrector improves the robustness and fault tolerance of the highly variable Wi-Fi signals. To fully reflect the performance of the proposed method, experiments have been carried out in the real environment of indoor parking lots. The results show that the proposed integrated model successfully provides reliable position estimation when the access points are lost under cybersecurity threats.

1. Introduction

In recent years, with the rapid development of computer science and mobile communication, and increasing market share of smart phones, tablet PCs, and other equipment, the need for location services has been growing in an unprecedented rapid pace. There are many potential applications emerging, such as real-time vehicle information service [1], traffic guidance information service, and parking guidance information service [2]. After years of development, Global Positioning System has been so maturely developed that sufficiently meets people's needs for outdoor positioning. However, the transmission of wireless signals by satellites and network base stations is inevitably obstructed by indoor structures, resulting in large signal deviations and failure of accurate positions in indoor positioning fields [3].

At present, varieties of technologies are continuously implemented in indoor positioning fields, including Ultrasonic Positioning [4, 5], Geomagnetic Positioning [6], Bluetooth Positioning [7, 8], UWB Positioning [9], and Wi-Fi Positioning [10, 11]. Among these technologies, Wi-Fi Positioning is the most prominent, as it has advantages in terms of vast communication range, low cost, convenient deployment, etc. [12, 13]. Meanwhile, almost all mobile terminals have built-in wireless network cards that can measure Wi-Fi signal strength, which can be used for indoor positioning [14, 15]. There are outstanding prospects for the development of Wi-Fi indoor positioning technology, and its research remains a meaningful and valuable work.

The Wi-Fi RSSI fingerprinting method is recognized as a main technology measure compared to geometric location method because it is flexible, is easy to identify, and does not need to obtain the physical location of the access points

[16, 17]. It consists of two phases: offline phase and online phase. In the offline phase, the fingerprint database with position label is constructed through the received signal strength indicator detected from various reference points and spatial coordinates of the reference points. In online phase, a location result is estimated through fingerprinting algorithm by comparing the received signal strength indicator collected from target points with the fingerprint database constructed in the previous phase.

The Wi-Fi RSSI fingerprinting method has a high practicality and effectiveness in most cases. However, when developing a Wi-Fi indoor positioning system in a real-world environment, there are several problems we have to face. With the large-scale deployment of Wi-Fi infrastructure, the cybersecurity threats of indoor positioning environments are getting more serious, leading to the fact that some access points' signal strength fluctuates extensively or even loses contact. In other words, the target point cannot obtain the received signal strength indicator that should have been received from some access points. This phenomenon directly leads to the invalidity of fingerprinting algorithm, and the indoor location system cannot get reliable application in a real-world environment.

To address this problem, researchers have thought of various approaches based on rules, hybrid, game theory, and graphs [18, 19], as well as discussing lots of anomaly detection models, e.g., models based on rules, clustering, vector supporting machine, closest proximity, and spectral decomposition [20, 21]. Zhang et al. [22] proposed an ellipse-type vector supporting machine to model the behavior attributes of the sensor data in wireless networks. However, ellipse-type support vector machine method has a problem on secondary optimization, which makes it impossible to be implemented on networks deployed in remote and harsh environments and is not suitable for general WLAN environments. Paola et al. [23] proposed an Adaptive Distributed Bayesian Approach for identifying outliers in collected data through wireless sensor networks, but the Bayesian Approach is not adaptive when offline, and its generalization practicability is mediocre. Mohammad Wazid [24] proposed a k-mean clustering to detect outliers and a mixed outlier detection method by obtaining parameter thresholds. Yenke et al. [25] proposed a distributed anomaly detection scheme based on Mahalanobis and Euclidean distance, which uses nearest neighbor search to improve the effectiveness of the algorithm. However, this method does not take into account the characteristics of access points, resulting in poor anomaly detection effect. On the other hand, these methods do not provide reliable location estimation methods after detecting an abnormality.

The emphasis of this work is to provide reliable position estimation when the access points' signal is lost under cybersecurity threats. Without knowing the physical location of the access points in the environment, we propose a new integrated model based on signal anomaly detector and signal distance corrector. In the offline phase, considering that Wi-Fi fingerprint database can be fitted into an n -dimensional surface in signal space, the signal anomaly detector is constructed based on signal

distortion theory and is trained through repeated comparison and analysis. In the online phase, for the unlabeled RSSI sample from the mobile terminal, the signal anomaly detector is used to realize online anomaly estimation, and the signal distance corrector is used to online distance correction. To fully reflect the performance of the proposed method, experiments are carried out in the real environment of indoor parking lots. The results show that the proposed integrated model yields higher anomaly detection accuracy and lower positioning mean error and makes it possible for application in cases when access points are lost under cybersecurity threats.

The rest of the paper is organized as follows. Section 2 introduces the related work of this paper, including Wi-Fi fingerprint data acquisition system, Wi-Fi real-time positioning system, and software architecture. Section 3 describes a new integrated model based on signal anomaly detector and signal distance corrector in detail. Section 4 presents the experimental design and results analysis. Section 5 summarizes the full paper and proposes suggestions for further research.

2. Related Work

2.1. Wi-Fi Fingerprint Data Acquisition System. The Wi-Fi fingerprint data acquisition system mainly completes the data acquisition of indoor Wi-Fi fingerprint signals in the offline phase and saves the collected Wi-Fi fingerprint signal data in the database as a data set for model training in subsequent work. The structure of Wi-Fi fingerprint data acquisition system is shown in Figure 1.

As shown in Figure 1, the fingerprint data acquisition system includes a client and a server. The client is an App based on the Android system, which mainly includes functions such as collecting Wi-Fi fingerprint data, uploading Wi-Fi fingerprint data, and displaying data upload results. The server includes a Web Server and a database. The Web Server responds to data collection requests, processes the data structure, and stores the processed data in the database. The database uses SQL Server to store massive Wi-Fi fingerprint data.

2.2. Wi-Fi Real-Time Positioning System. The Wi-Fi real-time positioning system mainly completes the data matching of indoor Wi-Fi fingerprint signals in the online phase. The entire system relies on the location fingerprint database collected in the offline phase and the online positioning model obtained in the offline training phase. The structure of Wi-Fi real-time positioning system is shown in Figure 2.

As shown in Figure 2, the real-time positioning system also includes a client and a server. The client is another App based on the Android system, which can scan the signals of Wi-Fi access points around the user in real time, send the scanned Wi-Fi information data to the server, and then receive the return from the server. As a result, the real-time location is displayed on the map. The server mainly uses Web Server to control the data flow jump and Python to achieve data preprocessing. The signal anomaly detector and

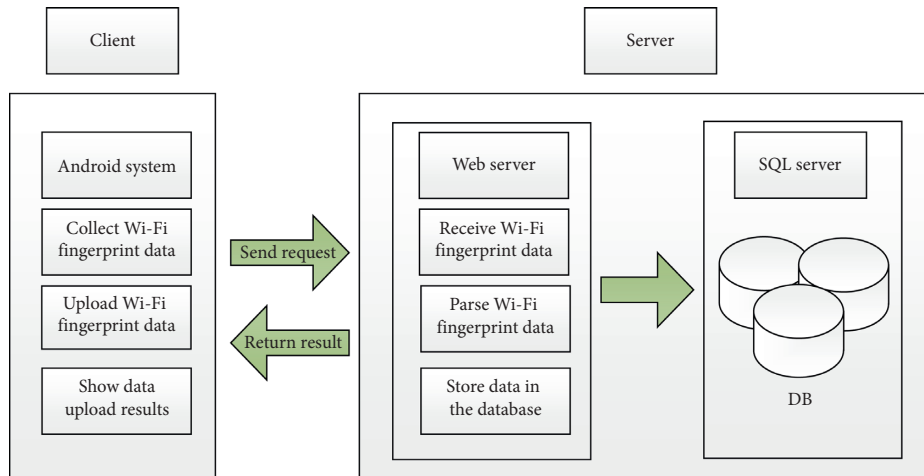


FIGURE 1: The structure of Wi-Fi fingerprint data acquisition system.

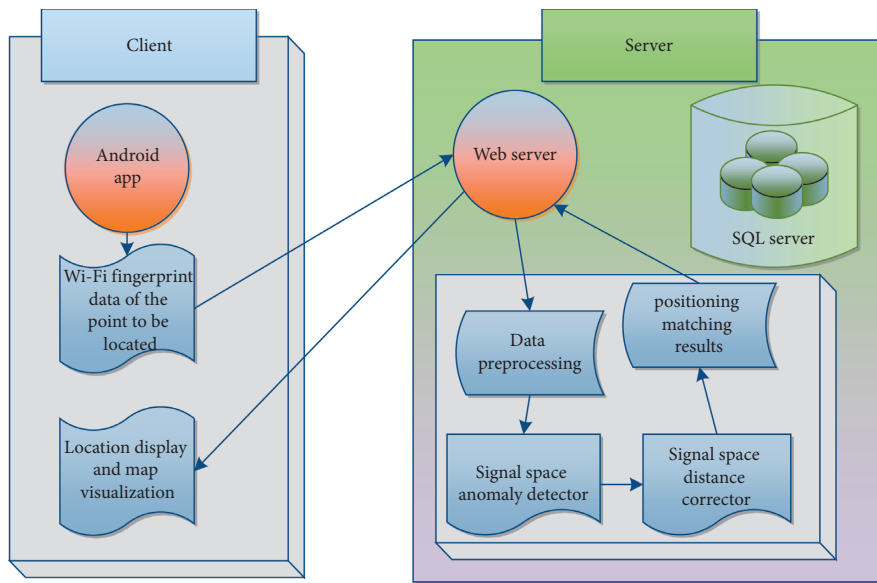


FIGURE 2: The structure of Wi-Fi real-time positioning system.

the anomaly distance corrector obtained through training are matched to obtain the current position coordinates, and the positioning result is returned to the client.

2.3. System Software. In this article, the system software is composed of client, server, and database. The client is based on Android system, the server is published on Tomcat 8.5, and the database uses Microsoft SQL server relational database. In addition, the proposed ensemble model is built by the integrated development environment of anaconda software based on Python3.6.

3. Proposed Methods

When some access points' signal strength fluctuates extensively or even loses contact due to the cybersecurity threats, the indoor location system cannot provide satisfied positioning accuracy in most situations. To address this

problem, we propose a new integrated model based on signal anomaly detector and signal distance corrector to provide reliable position estimation when the access points' signal is lost under cybersecurity threats. Figure 3 shows the process of data stream analysis. The process mainly contains two steps: the signal anomaly detector for online anomaly estimation and the signal distance corrector for online distance correct.

3.1. Signal Anomaly Detector. In this section, we will approximately describe the proposed signal anomaly detector. Firstly, we assume that the fingerprint data set of reference point can be expressed by $L = (r_1, l_1), (r_2, l_2), (r_3, l_3), \dots, (r_i, l_i), i = 1, 2, 3, \dots, n$, where n is the total number of APs. It is composed of the received signal strength vector of reference point $r_i = (r_{i1}, r_{i2}, \dots, r_{in})$ and the spatial coordinates of reference point $l_i = (x_i, y_i)$. Meanwhile, if the fingerprint data set of reference point can be fitted into an

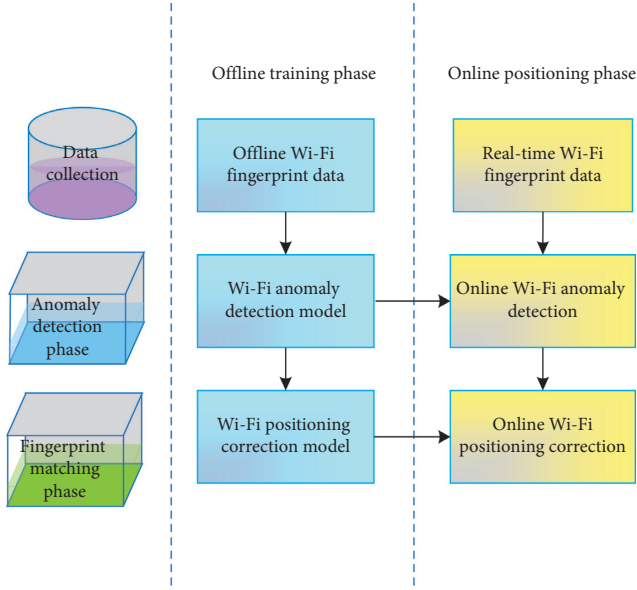


FIGURE 3: The block diagram of the new integrated model.

n -dimensional surface R , the (r_i, l_i) can be expressed as a point in n -dimensional signal space. If the received signal strength vector at the point to be located is expressed by $o = (o_1, o_2, \dots, o_n)$, the minimum distance D_{\min} from the point o to the n -dimensional surface R can be obtained by calculating the distance between o and r_i . Through the geometric rule theory, in the normal situation, we can conclude that D_{\min} must be able to converge to zero or a certain range, while, in the abnormal situation, D_{\min} will increase to a greater value than a certain range. Therefore, if D_{\min} cannot converge to a certain range, we can determine that it is in an abnormal situation. Figure 4 shows the sketch map of signal distortion.

The above signal anomaly detection method adopts Euclidean distance threshold to judge whether there is abnormality or not, which can be applied to anomaly detection in most cases [26, 27]. However, for Wi-Fi positioning system, the problem we must face is that the APs signal strength and signal stability are different for each point to be located because the positioning area is too large, and the environment is extremely complex. If we applied the above signal anomaly detection method to Wi-Fi positioning system, because it estimates Euclidean distance by all APs, it may exaggerate some small signal changes, and many normal signal changes may be judged as abnormal signal loss, resulting in it being difficult to accurately realize signal anomaly detection.

To solve this problem, we propose an improved signal anomaly detection algorithm by distinguishing trusted APs and untrusted APs. When estimating Euclidean distance in signal space, the proposed algorithm only considers the signal space distortion of trusted APs, while it does not consider the signal space distortion of untrusted APs, which narrows down some small signal changes and makes it easy to accurately judge whether there is an anomaly. The construction process of improved signal anomaly detection algorithm is discussed in detail as follows.

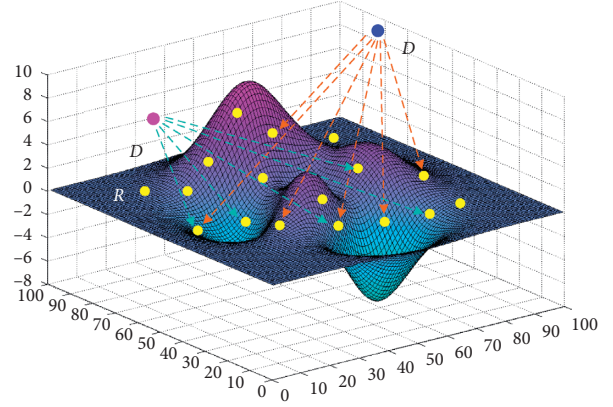


FIGURE 4: The sketch map of signal distortion.

Firstly, the AP set can be expressed by the $M = (AP_1, AP_2, \dots, AP_n)$, where n is the total number of APs. Meanwhile, the received AP signal strength at the point to be located can be expressed by $o = (o_1, o_2, \dots, o_n)$. As we know, the distance from each AP to the point to be located is different, so each AP signal strength at the point to be located is different. Meanwhile, because some APs are farther from the point to be located, it may not be able to search for some signals. These signals that cannot be searched may exaggerate some small signal changes, and many normal signal changes may be judged as abnormal signal loss. Therefore, we can define the APs corresponding to the signals that cannot be searched at the point to be located as A_{untrust} , and the APs corresponding to the signals that can be searched at the point to be located as A_{trust} .

Secondly, we can calculate the minimum Euclidean distance under A_{trust} and under A_{untrust} , respectively. The minimum Euclidean distance $D_{\min-\text{trust}}$ can be calculated by formula (1), and the minimum Euclidean distance $D_{\min-\text{untrust}}$ can be calculated by formula (2).

$$D_{\min-\text{trust}} = \sqrt{\sum_{AP_j \in A_{\text{trust}}} d_{ij}} = \sqrt{\sum_{AP_j \in A_{\text{trust}}} (r_{ij} - o_j)^2}, \quad (1)$$

$$D_{\min-\text{untrust}} = \sqrt{\sum_{AP_j \in A_{\text{untrust}}} d_{ij}} = \sqrt{\sum_{AP_j \in A_{\text{untrust}}} (r_{ij} - o_j)^2}. \quad (2)$$

Finally, according to the rules proposed above, when estimating Euclidean distance in signal space, we only consider the minimum Euclidean distance under A_{trust} , while we do not consider the minimum Euclidean distance under A_{untrust} . When $D_{\min-\text{trust}}$ is greater than distance threshold TH , it is judged that there is AP loss in o , which is represented in the abnormal situation, while when $D_{\min-\text{trust}}$ is less than distance threshold TH , it is judged that there is no AP loss in o , which is represented in the normal situation. Suppose that the abnormal state is defined as U , and the anomaly criterion can be expressed as follows:

$$U = \begin{cases} \text{true}, & D_{\min\text{-trust}} > TH \\ \text{false}, & D_{\min\text{-trust}} < TH \end{cases}. \quad (3)$$

3.2. Signal Distance Corrector. As mentioned above, the Wi-Fi fingerprinting method combines offline phase and online phase [28]. In the offline phase, the target is to realize the fingerprint data acquisition. The fingerprint data set of reference point can be expressed by (r_i, l_i) , which is composed of spatial coordinates at i -th the reference point $l_i = (x_i, y_i)$, $i = 1, 2, \dots, k$, where k is the total number of reference points, and received signal strength vector at the i -th reference point $r_i = (r_{i1}, r_{i2}, \dots, r_{in})$, where r_{ij} represents the RSSI signal of the j -th AP collected at the i -th reference point. In the online phase, the target is to estimate the location of the mobile terminals using the model. The received signal strength at the point to be located is $o = (o_1, o_2, \dots, o_n)$, and the position estimation l is obtained by calculating the similarity between o and r_i . Figure 5 shows the overall framework of the Wi-Fi fingerprinting method.

In addition, K-Nearest Neighbor (KNN) is usually used in the online phase [29]. KNN usually uses Euclidean distance to calculate similarity in signal space, and Euclidean distance can be calculated according to formula (4). The smaller the distance between them, the higher the similarity between them. Usually, the coordinates of the first K reference points are selected. The distance weighted K-Nearest Neighbor (DW-KNN) is based on the KNN [30], where the K nearest neighbors can be obtained by sorting according to formulae (5) and (6).

$$D_i = \sqrt{\sum_{j=1}^n (r_{ij} - o_j)^2}, \quad (4)$$

$$\hat{l} = \sum_{i=1}^k \omega_i l_i, \quad (5)$$

$$\omega_i = \frac{D^{-1}}{\sum_{j=1}^k D_j^{-1}}. \quad (6)$$

DW-KNN can be applied to Wi-Fi positioning estimation in most cases. However, for cybersecurity threats, the problem we must face is that the target point cannot obtain the received signal strength indicator that should have been received from some access points. When there is AP loss, if Euclidean distance is still calculated according to formula (4), the lost AP may make the K nearest neighbors obtained by sorting according to formulae (5) and (6) no longer reliable, leading to the invalidity of fingerprinting algorithm.

To address this problem, we propose a new fingerprint matching method based on DW-KNN. When the signal abnormality detector determines that there is AP loss, the proposed algorithm only considers the signal space distortion of trusted APs, while it does not consider the signal space distortion of untrusted APs, which greatly reduce the matching error caused by signal loss. Therefore, when the signal abnormality detector determines that there is AP loss, the corrected Euclidean distance can be calculated according to formula (7), while when the signal abnormality detector determines that there is no AP loss, Euclidean distance is still calculated according to formula (8). Suppose that the new Euclidean distance is defined as D , and the fingerprint matching criterion can be expressed as formula (9):

$$D_{\min\text{-abnormal}} = \sqrt{\sum_{AP_j \in A_{\text{trusted}}} d_{ij}} = \sqrt{\sum_{AP_j \in A_{\text{trusted}}} (r_{ij} - o_j)^2}, \quad (7)$$

$$D_{\min\text{-normal}} = \sqrt{\sum_{AP_j \in A_{\text{trust}}} d_{ij}} + \sqrt{\sum_{AP_j \in A_{\text{untrust}}} d_{ij}} = \sqrt{\sum_{AP_j \in A_{\text{trust}}} (r_{ij} - o_j)^2} + \sqrt{\sum_{AP_j \in A_{\text{untrust}}} (r_{ij} - o_j)^2}, \quad (8)$$

$$D = \begin{cases} D_{\min\text{-abnormal}}, & U = \text{true} \\ D_{\min\text{-normal}}, & U = \text{false} \end{cases}. \quad (9)$$

4. Experiments and Discussion

4.1. Experiment Environment. To evaluate the proposed strategy, we conducted a real experiment in the indoor parking lot of a shopping mall in China. This is a real indoor parking lot consisting of 150 parking spaces and covering an area of 2,000 square meters. Figure 6 is a real-life view of the indoor parking lot. The rectangular grid represents a parking space of 2.0 m \times 5.0 m. The signal identifier represents Wi-Fi access points, which has detected that 16 access points were

installed in the indoor parking lot. The real AP device for experiments is shown in Figure 7.

To enhance the robustness of the positioning system, three different types of mobile phones (Huawei glory 3C, Xiaomi mix2, and 360n5) are used in this paper. Meanwhile, data should be collected in all directions facing east, south, west, and north in turn to avoid the influence of the surrounding body on data collection. Then, we collected 100 training samples for four orientations by this way at every parking space. Meanwhile, to meet the scene change

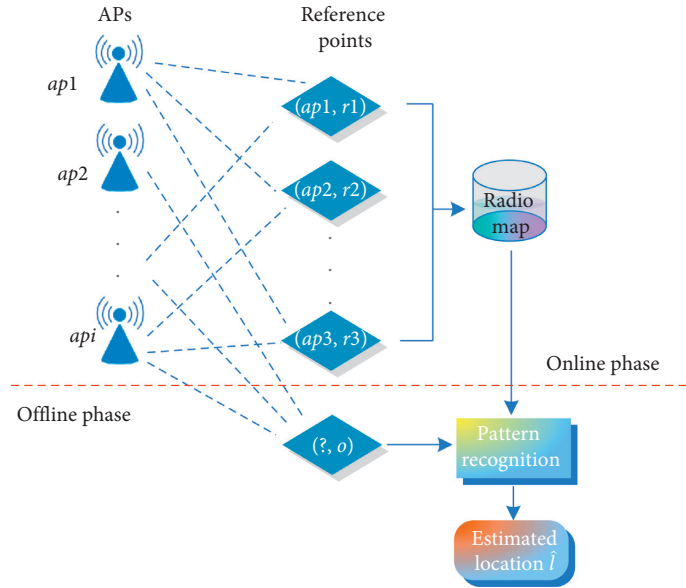


FIGURE 5: The overall framework of the Wi-Fi fingerprinting method.

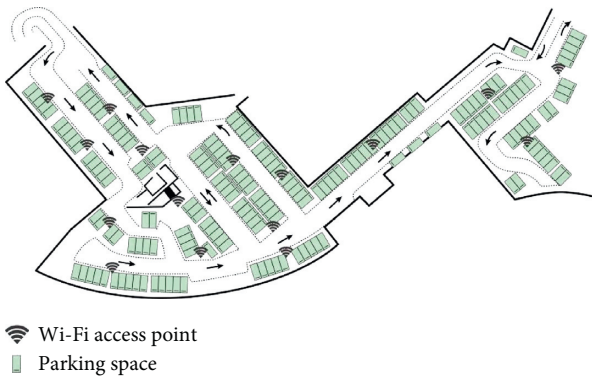


FIGURE 6: The real map of the indoor parking lot.



FIGURE 7: The real APs device for experiments.

problem in real application, we have taken training data and test data collected in different periods and collected 10 testing data pieces for four orientations by this way at every parking space after a week. As we know, because some APs are farther from the point to be located, it may not be able to search for some signals, and we record the missing value as -90 dB. The data storage structure is shown in Figure 8.

In order to compare the performances of various methods, three measuring standards of error distance,

precision, and accuracy are proposed. The error distance implies the Euclidean distance between the estimated coordinates and the true coordinates. The precision is another indicator of positioning performance, which is commonly described as the cumulative distribution function (CDF) of the error distance. The accuracy implies the average error distance of all pending points. The smaller the average error distance, the higher the accuracy, and vice versa.

4.2. The Necessity of the Proposed Integrated Model. In this section, to verify the necessity of the proposed signal anomaly detector and signal distance corrector, we adopt DW-KNN as the positioning classifier as it has the advantage of magnifying slight changes. As we know, we need to obtain the optimal number K for our experiments before constructing DW-KNN. Figure 9 is given the variation curve of the average positioning error with different number K , where the error bar represents the standard deviation. We can see that, with the continuous increase of the K value, the average positioning error value first gradually decreases and then slowly increases. It is pointed out that the average positioning error gets the minimum value when K equals seven, where the average positioning error is 2.59 m and the standard deviation is 1.91 m. It is equivalent to the fact that the DW-KNN positioning classifier has the best performance for Wi-Fi signal positioning when K equals seven.

On the basis of DW-KNN positioning classifier, we need to obtain the variation curve of the average positioning error in case of access points loss. To more intuitively show the variations of positioning accuracy in case of access points loss, we introduce the access points loss ratio $\alpha\%$ in each observation. Figure 10 shows the variation curve of the average positioning error with the change of the access points loss ratio. Meanwhile, the cumulative distribution function (CDF) of error distance with the change of the access points loss ratio is shown in Figure 11. We can see that

| Ap \ Parking lot | Ap | | | | | | | | | | | | | | | |
|------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 1000 | -90 | -53 | -54 | -66 | -62 | -66 | -84 | -85 | -82 | -73 | -83 | -82 | -90 | -90 | -90 | -79 |
| 1001 | -90 | -55 | -49 | -65 | -61 | -65 | -81 | -86 | -80 | -72 | -85 | -79 | -90 | -90 | -90 | -78 |
| 1002 | -90 | -55 | -45 | -64 | -59 | -65 | -80 | -83 | -79 | -71 | -83 | -77 | -90 | -90 | -90 | -81 |
| 1003 | -90 | -59 | -50 | -62 | -61 | -64 | -79 | -84 | -79 | -75 | -82 | -81 | -90 | -90 | -90 | -83 |
| 1004 | -90 | -59 | -47 | -60 | -58 | -64 | -80 | -84 | -80 | -77 | -85 | -80 | -90 | -90 | -90 | -81 |
| 1005 | -90 | -63 | -48 | -58 | -57 | -60 | -77 | -80 | -79 | -78 | -83 | -80 | -90 | -90 | -90 | -80 |
| 1006 | -90 | -63 | -51 | -56 | -55 | -54 | -74 | -80 | -76 | -76 | -80 | -80 | -90 | -90 | -90 | -78 |
| 1007 | -90 | -67 | -53 | -56 | -54 | -55 | -72 | -80 | -74 | -76 | -79 | -78 | -90 | -90 | -90 | -76 |
| 1008 | -90 | -68 | -54 | -53 | -54 | -58 | -75 | -81 | -76 | -77 | -80 | -78 | -90 | -90 | -90 | -77 |
| 1009 | -90 | -68 | -53 | -49 | -56 | -52 | -72 | -79 | -76 | -74 | -78 | -77 | -90 | -90 | -90 | -77 |
| 10010 | -90 | -68 | -55 | -44 | -54 | -51 | -70 | -79 | -75 | -71 | -77 | -78 | -90 | -90 | -90 | -74 |
| 10011 | -90 | -70 | -58 | -49 | -58 | -53 | -71 | -78 | -73 | -76 | -76 | -77 | -90 | -90 | -90 | -78 |
| 10012 | -90 | -70 | -58 | -51 | -60 | -54 | -67 | -76 | -70 | -76 | -75 | -78 | -90 | -90 | -90 | -76 |
| 10013 | -90 | -53 | -56 | -63 | -43 | -64 | -79 | -82 | -80 | -73 | -83 | -87 | -90 | -90 | -90 | -77 |
| 10014 | -90 | -61 | -58 | -58 | -47 | -58 | -80 | -80 | -80 | -72 | -79 | -85 | -90 | -90 | -90 | -76 |

FIGURE 8: The data storage structure.

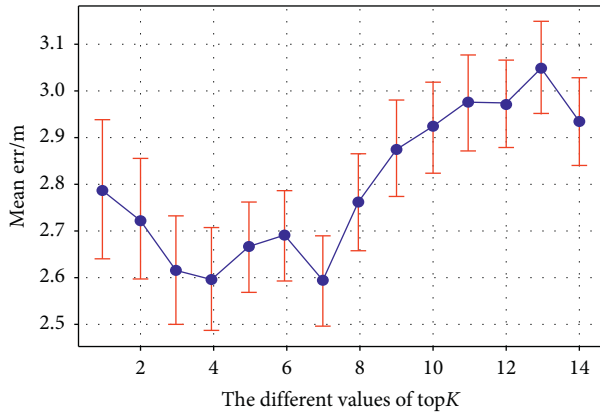


FIGURE 9: The variation curve of the average positioning error with different number K .

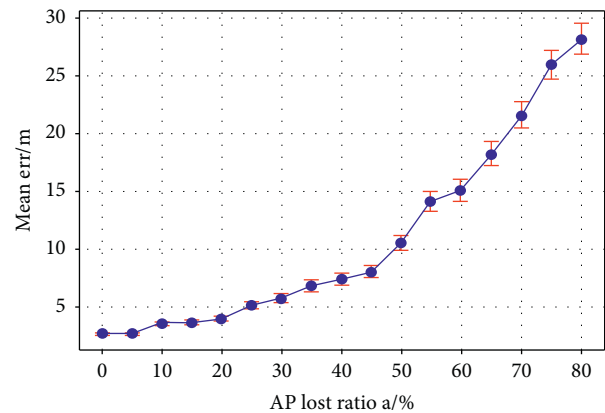


FIGURE 10: The variation curve of the average positioning error with the change of the access points loss ratio.

the average positioning error continues to grow as the access points loss ratio continues to increase. When the access points loss ratio $\alpha = 10\%$, the average positioning error increases to 4.16 m, which is an increase of 31.7% compared to it in nonabnormal conditions, while, for all observations, the access points loss ratio 10% is only one or two access points, where there are up to sixteen access points in all observations. When the access points loss ratio $\alpha = 25\%$, the average positioning error increases to 5.13 m, which is an increase of 98.1% compared to it in nonabnormal conditions. It is worth mentioning that when the access points loss ratio $\alpha = 55\%$, the average positioning error is as high as 14.06 m. It can be obtained that the proposed signal anomaly detector and signal distance corrector proposed are indispensable.

4.3. Construction of Signal Anomaly Detector. As mentioned in section 3, the construction of the signal anomaly detector needs to determine the distance threshold. Meanwhile, the distance threshold can be obtained by the cumulative distribution function of D_{\min} . The positive detection rate and false detection rate decrease with the increase of the distance threshold; nevertheless, it is the decrease of the distance threshold instead of the law. The goal of distance threshold selection is to satisfy the balance between positive detection rate and false detection rate. To verify that the proposed new signal anomaly detector can achieve better anomaly detection performance, we adopt a traditional anomaly detection method that does not distinguish between trusted APs and untrusted APs for comparative analysis.

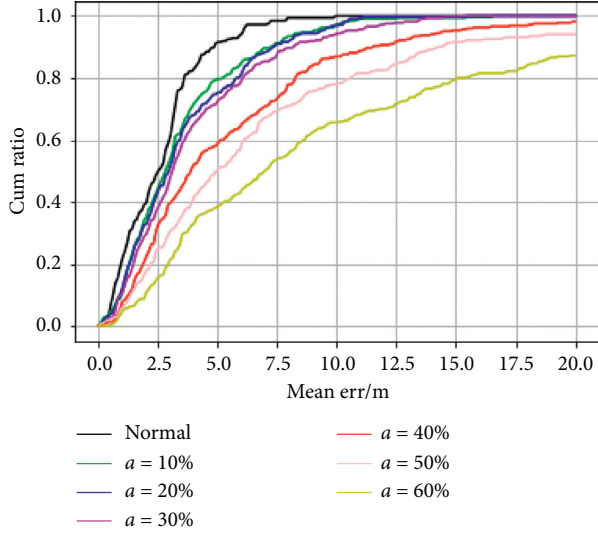


FIGURE 11: The cumulative distribution function (CDF) of error distance with the change of the access points loss ratio.

Figure 12 shows the cumulative distribution function of D_{\min} with the change of the access points loss ratio in the traditional anomaly detection method. Meanwhile, the cumulative distribution function (CDF) of D_{\min} with the change of the access points loss ratio in the improved signal anomaly detector is shown in Figure 13. We can see that when the access points are lost, both curves move to the right, and the greater the ratio of access points loss, the more obvious the right shift. But no matter what ratio of AP is lost, the improved signal anomaly detector moves to the right obviously compared to the traditional anomaly detection method. This phenomenon indicates that the improved signal anomaly detector is more sensitive.

As shown in Figure 12, the best distance threshold is 20 dBm by analyzing the curve in the traditional anomaly detection method. Therefore, when 10%, 20%, 30%, 40%, 50%, and 60% of the APs are lost, there will be 40%, 42%, 69%, 76%, 83%, and 88% of the positive detection rate, and the corresponding false detection rate is 4%. When the AP loss ratio is less than 30%, this traditional signal anomaly detector is not in high implementability, so it is obliged to improve the traditional anomaly detection method.

As is shown in Figure 13, the best distance threshold is 14 dBm by analyzing the curve in the improved signal anomaly detector. Therefore, when 10%, 20%, 30%, 40%, 50%, and 60% of the access points are lost, there will be 60%, 64%, 82%, 85%, 93%, and 95% of the positive detection rate, and the corresponding false detection rate is 1%. When the AP loss ratio is less than 30%, compared with the traditional anomaly detection method, it has increased by 47%, 49%, and 19%, respectively, and the false detection rate of the system is also significantly reduced. It can verify that the proposed new signal anomaly detector can achieve reliable anomaly detection performance.

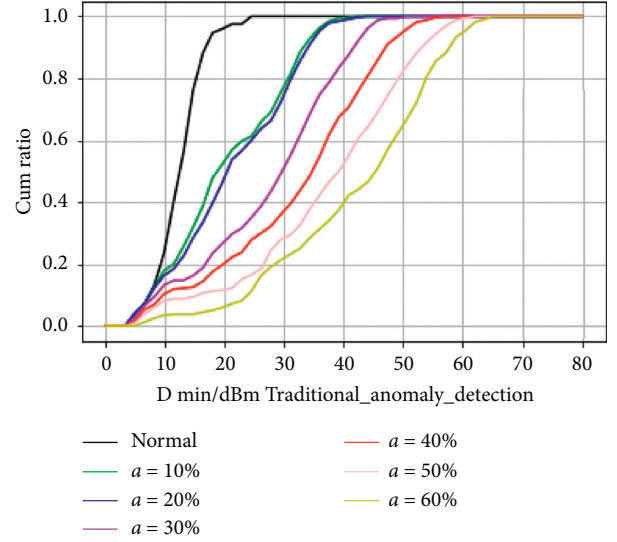


FIGURE 12: The cumulative distribution function (CDF) of D_{\min} with the change of the access points loss ratio in the traditional signal anomaly detector.

4.4. Construction of Signal Distance Corrector. On the basis of the signal anomaly detector in the previous step, we propose a new fingerprint matching method based on DW-KNN for abnormal situation. When the signal abnormality detector determines that there is AP loss, the proposed algorithm only considers the signal space distortion of trusted APs, while it does not consider the signal space distortion of untrusted APs, which greatly reduce the matching error caused by signal loss.

Figure 14 shows the variation curve of the average positioning error with the change of the access points loss ratio. Meanwhile, the cumulative distribution function (CDF) of error distance with the change of the access points loss ratio is shown in Figure 15. When the access points loss ratio $\alpha = 10\%$, the average positioning mean error increases to 3.12 m, which is an increase of 20.4% compared to it in normal situation. When the access points loss ratio $\alpha = 25\%$, the average positioning mean error increases to 3.62 m, which is an increase of 39.7% compared to it in normal situation. It is worth mentioning that when the access points loss ratio $\alpha = 55\%$, the average positioning error grows gradually.

To verify that the proposed signal distance corrector can achieve better positioning performance under abnormal situation, traditional DW-KNN algorithm is used to compare the performance of the algorithm. The results of the experiment on the positioning mean error of the traditional DW-KNN algorithm and the improved signal distance corrector with the change of the access points loss ratio are shown in Figure 16. We can see that the average error increase of the improved DW-KNN was significantly gentler than that of the traditional DW-KNN. Meanwhile, if the acceptable average positioning error is within 5.00 m, the traditional DW-KNN algorithm can tolerate 25% AP loss in

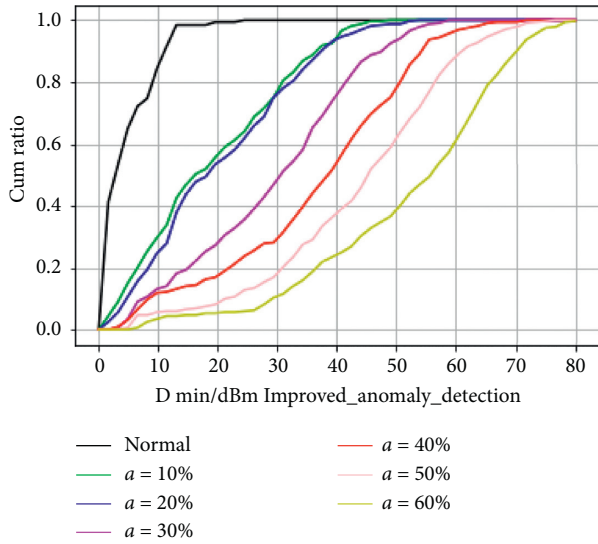


FIGURE 13: The cumulative distribution function (CDF) of D_{\min} with the change of the access points loss ratio in the improved signal space anomaly detector.

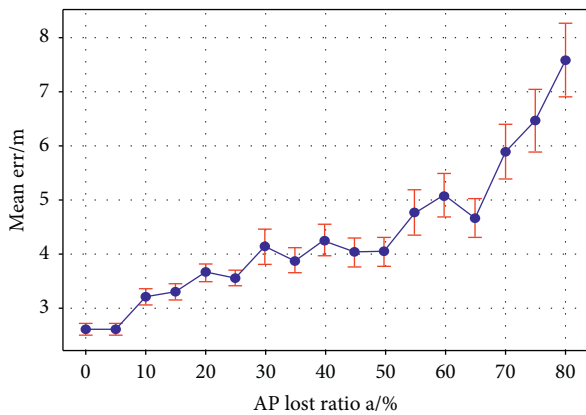


FIGURE 14: The variation curve of the average positioning error with the change of the access points loss ratio.

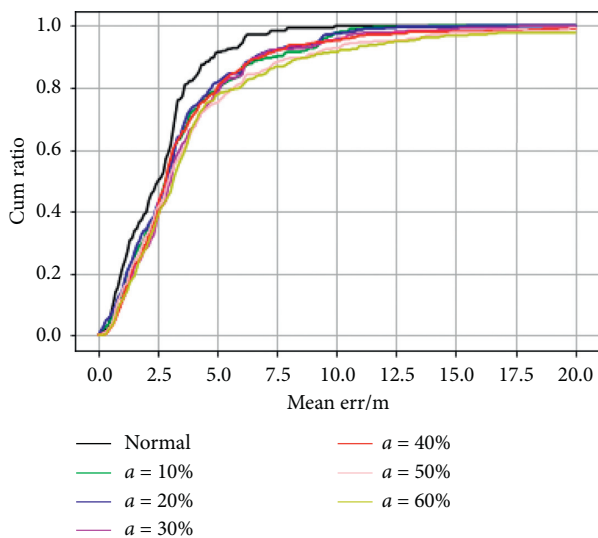


FIGURE 15: The cumulative distribution function (CDF) of error distance with the change of the access points loss ratio.

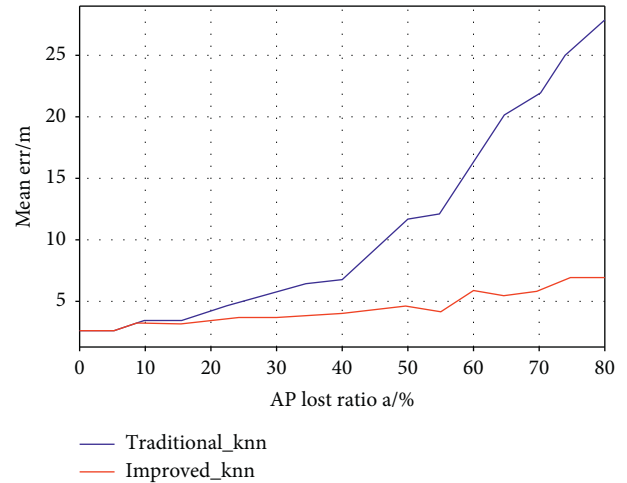


FIGURE 16: The positioning mean error of the traditional DW-KNN algorithm and the improved signal distance corrector with the change of the access points loss ratio.

the observations, while the improved signal distance corrector can tolerate approximately 60% and shows extremely strong robustness to access point loss, which proves that the improved signal distance corrector is more reliable under abnormal conditions. It is worth mentioning that when the access points loss ratio $\alpha = 80\%$, the improved DW-KNN can improve the average error by 20 m at most or so compared with the traditional DW-KNN. It can verify that the proposed new signal distance corrector can achieve reliable distance correction performance.

5. Conclusion and Future Research

In this paper, we propose a new integrated model based on signal anomaly detector and signal distance corrector. The signal anomaly detector improves recognition capability of the uncertain signal and noise, while the signal distance corrector improves the robustness and fault tolerance of the highly variable Wi-Fi signals. In the offline phase, considering that Wi-Fi fingerprint database can be fitted into an n-dimensional surface in signal space, the signal anomaly detector is constructed based on signal distortion and is trained through repeated comparison and analysis. In the online phase, for the unlabeled RSSI sample from the mobile terminal, the signal anomaly detector is used to realize online anomaly estimation, and the signal distance corrector is used to online distance correct. To fully reflect the performance of the proposed method, experiments have been carried out in the real environment of indoor parking lots. The results show that the proposed integrated model successfully provides reliable position estimation when the access points signal is lost under cybersecurity threats.

In the future, we are planning to solve the problem of best-discriminating AP optimization in large-scale complex environments with partition walls. In addition, we plan to integrate other mobile phone sensors (such as Bluetooth and Geomagnetism) to obtain better positioning precision.

Data Availability

The data are true and reliable, and the original data have been saved in the attachment.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

H. W. conceptualized the study. Z. Y. performed data curation. Z. Y. performed investigation. Z. Y. wrote the original draft. Y. C., Z. C., and X. Z. reviewed & edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This work was supported by Natural Science Foundation of China (NSFC) under Grant nos. 62073250 and 62003249, in part by Key Research and Development Program of China under Grant 2017YFC0806503-05, in part by Key Research and Development Program of Hubei Province under Grant 2020BAB021, and in part by Science and Technology Research Project of Hubei Provincial Department of Education under Grant D20201105.

References

- [1] S.-H. Jung, G. Lee, and D. Han, "Methods and tools to construct a global indoor positioning system," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 6, pp. 906–919, 2018.
- [2] L. Ma, Z. You, T. Liu, and S. Shi, "Coupled integration of CSAC, MIMU, and GNSS for improved PNT performance," *Sensors*, vol. 16, no. 5, p. 682, 2016.
- [3] A. Brack, "Erratum to: reliable GPS+BDS RTK positioning with partial ambiguity resolution," *Gps Solutions*, vol. 21, no. 3, pp. 1–10, 2017.
- [4] Y. Shi and C. Zhao, "Simple new ultrasonic piezoelectric actuator for precision linear positioning," *Journal of Electroceramics*, vol. 28, no. 4, pp. 233–239, 2012.
- [5] J. F. Huang, C. C. Liu, and K. F. Lin, "Enhancing ultrasonic robot positioning accuracy with parallel codes acquisition of composite pseudo-noise sequences," *Neuroimage*, vol. 60, no. 1, pp. 340–352, 2015.
- [6] Y. Keiko, O. Shintaro, O. Tomohiro et al., "Smart hospital infrastructure: geomagnetic in-hospital medical worker tracking," *Journal of the American Medical Informatics Association*, vol. 28, no. 3, pp. 477–486, 2021.
- [7] C. Y. T. Kwok, M. S. Wong, S. Griffiths et al., "Performance evaluation of iBeacon deployment for location-based services in physical learning spaces," *Applied Sciences*, vol. 10, no. 20, p. 7126, 2020.
- [8] F. J. Aranda, F. Parralejo, F. J. Álvarez, and J. Torres-Sospedra, "Multi-slot BLE raw database for accurate positioning in mixed indoor/outdoor environments," *Data*, vol. 5, no. 3, p. 67, 2020.
- [9] D. T. A. Nguyen, H.-G. Lee, E.-R. Jeong, H. L. Lee, and J. Joung, "Deep learning-based localization for UWB systems," *Electronics*, vol. 9, no. 10, p. 1712, 2020.
- [10] C. Luo, L. Cheng, M. C. Chan, Y. Gu, J. Li, and Z. Ming, "Pallas: self-bootstrapping fine-grained passive indoor localization using WiFi monitors," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 466–481, 2017.
- [11] Z. Yuan, X. Zha, and X. Zhang, "Adaptive multi-type fingerprint indoor positioning and localization method based on multi-task learning and weight coefficients K-nearest neighbor," *Sensors*, vol. 20, no. 18, p. 5416, 2020.
- [12] C.-C. Huang and H.-N. Manh, "RSS-based indoor positioning based on multi-dimensional kernel modeling and weighted average tracking," *IEEE Sensors Journal*, vol. 16, no. 9, pp. 3231–3245, 2016.
- [13] A. Khalajmehrabadi, N. Gatsis, and D. Akopian, "Modern WLAN fingerprinting indoor positioning methods and deployment challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1974–2002, 2017.
- [14] D. Yu, Q. Hu, and S. Wang, "A drift-of-stay pattern extraction method for indoor pedestrian trajectories for the error and accuracy assessment of indoor wi-fi positioning," *ISPRS International Journal of Geo-Information*, vol. 8, no. 11, p. 468, 2019.
- [15] J. Chen, Y. Zhang, and W. Xue, "Unsupervised indoor localization based on smartphone sensors, iBeacon and wi-fi," *Sensors*, vol. 18, no. 5, p. 1378, 2018.
- [16] L. Zhang, S. Valaee, Y. Xu, L. Ma, and F. Vedadi, "Graph-based semi-supervised learning for indoor localization using crowdsourced data," *Applied Sciences*, vol. 7, no. 5, p. 467, 2017.
- [17] S. He, S.-H. Gary Chan, "Wi-fi fingerprint-based indoor positioning: recent advances and comparisons," *IEEE Communication Surveys and Tutorials*, vol. 18, no. 1, pp. 466–490, 2016.
- [18] X. Xu, "Sequential anomaly detection based on temporal-difference learning: principles, models and case studies," *Applied Soft Computing*, vol. 10, no. 3, pp. 859–867, 2010.
- [19] A. Wu and Z. Zeng, "Global Mittag-Leffler stabilization of fractional-order memristive neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 1, pp. 206–217, 2017.
- [20] A. Wu, H. Liu, and Z. Zeng, "Observer design and H_∞ performance for discrete-time uncertain fuzzy-logic systems," *IEEE Transactions on Cybernetics*, 2020, In press.
- [21] J. Yang, C. Zhou, S. Yang et al., "Anomaly detection based on zone partition for security protection of industrial cyber-physical systems," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 5, 2017.
- [22] Y. Zhang, N. Meratnia, and P. J. M. Havinga, "Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine," *Ad Hoc Networks*, vol. 11, no. 3, pp. 1062–1074, 2013.
- [23] A. De Paola, S. Gaglio, G. L. Re, F. Milazzo, and M. Ortolani, "Adaptive distributed outlier detection for WSNs," *IEEE Transactions on Cybernetics*, vol. 45, no. 5, pp. 902–913, 2015.
- [24] M. Wazid, "Hybrid anomaly detection using K-means clustering in wireless sensor networks," *IACR Cryptology ePrint Archives*, vol. 2014, p. 712, 2014.
- [25] B. O. Yenke, M. Aboubakar, C. Titouna et al., "Adaptive scheme for outliers detection in wireless sensor networks," *International Journal of Computer Networks and Communications Security*, vol. 5, no. 5, p. 105, 2017.
- [26] K. I. Ranney and M. Soumekh, "Hyperspectral anomaly detection within the signal subspace," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 3, pp. 312–316, 2006.

- [27] S. Maeda and H. Shibuya, “Anomaly detection method and anomaly detection system,” *U.S. Patent Application*, vol. 13/144, no. 343[P], pp. 2–16, 2012.
- [28] F. M. Lopez-Rodriguez and F. Cuesta, “An android and arduino based low-cost educational robot with applied intelligent control and machine learning,” *Applied Sciences*, vol. 11, p. 48, 2021.
- [29] M. Mejdoub and C. Ben Amar, “Classification improvement of local feature vectors over the KNN algorithm,” *Multimedia Tools and Applications*, vol. 64, no. 1, pp. 197–218, 2013.
- [30] J. Aguilera, L. C. González, M. Montes-y-Gómez et al., “A new weighted K-nearest neighbor algorithm based on Newton’s gravitational force,” in *Proceedings of Iberoamerican Congress on Pattern Recognition*, pp. 305–313, Springer, Madrid, Spain, November 2018.