A Fingerprint Location Framework for Uneven WiFi Signals Based on Machine Learning

Xu Lu, Kejie Zhong, Zhiwei Guan, and Jun Liu

***Abstract*—WiFi fingerprint positioning is a common method for indoor location determination. Existing methods are susceptible to fluctuations in WiFi signal strength during the offline phase, leading to unevenly received signals. Additionally, during online positioning, there is a lack of integration with historical trajectory information. These problems can result in errors in both offline fingerprint acquisition and online location positioning. To address these problems, we propose a method that combines normality detection in the offline phase and Location Weighted K-nearest Neighbor(LWKNN) positioning in the online phase. In the offline phase, initial Received Signal Strength Indication(RSSI) samples undergo preprocessing based on skewness and kurtosis for normality detection. If the samples conform to a normal distribution model, the probability density is estimated using the normal distribution function. If not, estimation occurs using the kernel density function model. Subsequently, values are averaged after Kalman filtering to establish a high-precision fingerprint database. During the online positioning phase, the LWKNN algorithm is employed. Initially, the Weighted K-nearest Neighbor method estimates the position, and this information is utilized as features to construct a Long Short Term Memory(LSTM) network model. The optimal path is determined through the least square method. Finally, the obtained outputs are integrated with historical data from the fingerprint positioning trajectory to enhance target positioning accuracy. Experimental results demonstrate that our indoor localization method significantly improves WiFi fingerprint localization accuracy compared to traditional localization methods.**

Link to graphical and video abstracts, and to code: https://latamt.ieeer9.org/index.php/transactions/article/view/8668

***Index Terms*—WiFi fingerprint positioning, Normality detection, Trajectory prediction, LWKNN, LSTM.**

I. Introduction

The Internet of Medical Things [1] (IoMT) was born as the number of connected medical devices increased, supporting advances in medical level data collection and transmission, interconnection technologies, service systems, and software. In terms of smart patient services, IoMT is able to target special or critically ill patients in the hospital, and services include inhospital navigation, personnel location, and alarm assistance. By wearing a smart bracelet on the patient and using its indoor positioning function, medical staff can view the patient's walking route and realtime location in real time according to the intelligent monitoring system, and once the patient goes out of the limited area, an alarm can be generated, and then medical staff will assist the patient to return to the safe area in time to prevent accidents. In addition, when the patient is in discomfort and urgently needs help, a key alarm can be realized through the smart wearable device [2], and medical personnel can respond quickly to protect the patient's life safety.

Therefore, for indoor places with complex environments such as hospitals and nursing homes, it is important to locate the location of the target effectively and precisely in a timely manner [3]. For this problem, many technical solutions have been explored to meet the demand for indoor location services. Currently, WiFi is widely used for public networks in various cities, as well as home and office networks [4]. Therefore, WiFi-based indoor positioning technology can make full use of ubiquitous WiFi signals without requiring any additional hardware equipment, greatly reducing the cost positioning while ensuring positioning accuracy and high signal coverage [5]. The most widely used algorithm for WiFi-based indoor positioning technology is the fingerprint positioning method [6]. The method consists of two stages [7]: The training stage (Offline Stage) and the online positioning stage. In the training stage, the main task is to evenly and reasonably set several fingerprint reference points across the positioning area. Then, we collect the signal at the location of the fingerprint reference point (Pedestrian), and use it to establish the fingerprint. In the online positioning stage, the goal is to collect the Access Point signal value at the target location in real time, then estimate the exact location of the point with the location fingerprint positioning algorithm.

WiFi signals are easily detected in indoor environments such as hospital, but they are also easily affected by the outside world during propagation. External factors, such as walls, the ground, human bodies, temperature, and humidity, can reflect and scatter WiFi signals during propagation. This causes high variation with respect to time in the received RSSI signal at a fixed position in the room [8]. The most common methods to deal with this problem are mean model [9], median model and normal distribution model [10]. The normal distribution model uses the normal distribution function to filter out signal values with high probability, which results in a better signal filtering effect than the mean and median models. However, in complex indoor environments, not all samples satisfy the normal distribution, so the normal distribution model does not provide accurate estimates for the total number of RSSI signal samples received by each fingerprint point.

Xu Lu is with the Guangdong Polytechnic Normal University, Guangdong Provincial Key Laboratory of Intellectual Property & Big Data, and Pazhou Lab, Guangzhou, China (e-mail: bruda@126.com).

Kejie Zhong, Zhiwei Guan and Jun Liu are with the School of Computer Science, Guangdong Polytechnic Normal University, Guangdong, China (e-mails: 1043485159@qq.com, kk18637199812@163.com and liujun7700@163.com).

For the online positioning stage, a Nearest Neighbor algorithm can locate the position of the target, but struggles to predict the category if sample numbers are unbalanced. Thus, the accurate position of the target cannot be accurately determined. The K-nearest Neighbors method finds the K data vectors nearest to the measured vector, then averages the position coordinates of the K vectors to estimate the position output of the terminal. However, in the case of the same number of AP points, the weight of AP points that dominate the target location positioning is not considered. Each AP point under the same weight will bring greater positioning error. The Weighted K-nearest Neighbors method [11] differs from the KNN algorithm by multiplying the coordinates of each database vector by a weighting coefficient. However, this method does not consider that the signal strength similarity is not completely consistent with the position similarity in the online matching stage, which amplifies the positioning error.

In order to solve the above two problems, this paper proposes a positioning method that combining skewness and the kurtosis normality test [12] with the trajectory prediction model. In the offline phase, we detect the overall distribution of RSSI signal samples by skewness kurtosis. If the sample population accepts the null hypothesis, we use the normal distribution method to estimate the probability density of the fingerprint points. If the normal model is not satisfied, we instead use the kernel density function to estimate the probability density, then filter out high probability signals. After smoothing the RSSI data, we store the average value in the database. In the online positioning stage, we first locate the target using WKNN, then establish an LSTM model using the obtained location information as the feature information. We use the optimized path obtained by the least square method as the output to supervise the training. Finally, we obtain more accurate target positioning by combining our output with the historical data of fingerprint positioning trajectory. The main contributions of this paper include:

a) A positioning method based on normality detection and trajectory prediction, which is a fusion of offline and online positioning methods, that can adaptively adjust the positioning process.
b) The normality detection module proposed in this paper divides and preprocesses the fingerprint information in the existing fingerprint database without re-collecting signal data.
c) The trajectory prediction model is combined with fingerprint positioning for the first time. This method can effectively utilize positioning history information.

II. Related Work

In recent years, researchers have made many improvements in the acquisition and application of WiFi signals in the offline phase. Luo et al. [13] proposed a Gaussian Mixture Model (GMM) for offline fingerprint clustering. The positioning area is divided into several subregions through offline data, and the corresponding subregion label training forest-based random subregion classifier. This method can significantly reduce errors without any hardware calibration. To address the problems of signal instability and fingerprint drift, Wang et al. [14] proposed a preprocessing method of RSSI value and CSI amplitude value based on Kalman filters and the Gaussian function, and added an improved CSI phase after a linear transformation. After effectively eliminating the mutation and noise data, one can achieve accurate and smooth outputs of RSSI and CSI values. Finally, they perform dimensionality reduction on the obtained high dimensional data values and establish a fingerprint database. The method performs well on tasks such as denoising, fusion positioning, and realtime filtering. Guo et al. [15] proposed a method to construct multi-fingerprint groups by collecting Hyperbolic position Fingerprints and signal intensity Difference Fingerprints from RSS fingerprints. The offline phase obtains MFTC by obtaining multiple fingerprint groups and continuous training of each basic classifier for each fingerprint. This method can enhance the RSSI fingerprint, but does not consider the impact of WIFI timevarying. Huang et al. [16] proposed a new scheme to adapt radio maps to environmental dynamics online by using lowcost crowdsourcing RSS measurement. A coarse grained radio map is created by using standard Gaussian Process Regression in the offline phase. Extended GPR alleviates drops in model accuracy caused by noise location labels, and can effectively improve localization and positioning accuracy. One disadvantage of this approach is that the radio map needs to be updated several times in real time. Li et al. [17] proposed a localization method based on sparse fingerprint acquisition and an improved weighted K nearest neighbor algorithm. The reference points are sparsely selected, and the abnormal values of the collected RSS are preprocessed according to the faulttolerant quartile method. The Gaussian process regression model is trained on the processed fingerprint data. Optimal hyperparameters for the model are obtained by a symbiotic biological search algorithm, which improves the generalization ability of the model. Finally, they predict the RSS of the nonreference points in the positioning area, and the database is built. This method has good strong fingerprint prediction ability and positioning performance, but does not consider the impact of time varying RSS signals in the actual place. Thus, it is only suitable for indoor static positioning.

Researchers have also made many improvements in the online positioning stage by further developing the weighted K nearest neighbor method. Kong et al. [18] divided the reference points into K subclasses with the K-means clustering algorithm, so as to reduce the fingerprint search space and improve matching efficiency. They then combine the weighted calculation of the adjacent points in the selected subclass with WKNN to improve the calculation proportion of the higher correlation reference points. Cao et al. [19] first analyzed the geometric structure of the K-nearest point, then removed the point with the longest distance between the center of the K-nearest point and itself. They proceed to analyze the geometric position of the unknown point, and the weights of the algorithm are determined by the geometric distance between the adjacent point and the center point, as well as the Euclidean distance between the adjacent point and the unknown point. This method avoids the use of matching points with deviations in positioning calculations, and improves stability and accuracy of positioning. However, it does not consider the positional stability at the actual boundary, and as a result, the positioning error at the geometric edge of the fingerprint database is still large. Chen et al. [20] proposed a variance-based weighted distance to improve the WKNN algorithm. The calculation weight of the distance is designed according to the variance information of the signal strength distribution of the sampling points. The similarity of the unit point is calculated by the weighted distance to weaken the influence of the RSS unstable access point. This method effectively reduces the influence of WiFi signal instability on positioning effect and improves positioning accuracy. However, the actual fingerprint information is affected by various factors in the actual situation. Collected data has a large matching error with the fingerprint database, which is only suitable for static environment. Pan et al. [20] proposed an AHPWKNN indoor positioning method combining AHP technology and weighted K nearest neighbor algorithm. They use AHP to assign weights when using WKNN to select fingerprints for positioning. AHP technology amplifies the influence of the received signal strength gap between reference points on the weight, which leads to better positioning performance. This method is robust against fluctuations in the RSSI, as well as deviations from the measured RSS. Zhao et al. [21] proposed an asymmetric Gaussian filtering algorithm IWKNN based on the signal strength distribution characteristics of smart venues. The method combines a specific signal distribution model and proposes asymmetric Gaussian filtering to improve generalization ability. Because it meets the requirements of realtime and high precision, it obtains lower delay by discarding useless information and improves the utilization of the database to ensure higher precision. Li et al. [22] proposed a K-nearest neighbor indoor fingerprint localization method based on a coarse localization circle and a highest similarity threshold. This method forms a circular domain in the coarse positioning process, reduces the positioning range, and solves the interference problem of irrelevant fingerprints. Additionally, they introduce the faulttolerant mechanism to dynamically adjust the circle domain to ensure that the coarse positioning circle domain contains the highest similarity reference point, which improves the fault-tolerant ability of coarse positioning.

We propose an indoor positioning algorithm that integrates skewness and kurtosis for normality detection and trajectory prediction in order to effectively reduce the influence of diffraction, scattering, reflection and timevarying of WiFi signals during propagation. The influence of WiFi signal instability arises in two different forms. Unstable RSSI measurements during fingerprint acquisition leads to the inability to establish an accurate fingerprint database. Additionally, unstable RSSI values in the positioning stage greatly affect the final result of the WKNN matching algorithm.

The proposed method not only preprocesses the collected fingerprint information to establish a high precision WiFi fingerprint database, but also makes effective use of localization history information to reduce the error of the matching algorithm caused by WiFi time variation. Compared with traditional localization methods, our proposed algorithm obtions higher localization accuracy.

III. Method

The general framework of the proposed method is shown in Fig. 1.

The offline stage analyzes the received signal strength values by skewness and kurtosis analysis. The collected signal values can be processed in the following two cases. The first case is that the collected signal values follow a normal distribution (black in the figure), where the probability density of the sample is estimated by using a normal distribution function. The second case is that a small number of sample values do not follow a normal distribution (blue in the figure), so we instead use the kernel density function to calculate the overall probability of the sample. Our method also uses the RSSI value processed by the Kalman filter algorithm to smooth the RSSI data. Finally, we record the average value as the determined RSSI value into the fingerprint database. In the online positioning stage, we first perform preliminary position estimation through the WKNN algorithm. Then, we use historical positioning information as feature information to establish the LSTM model. Through training, predicted results are fed back to the next target location positioning, which excludes location points with large deviation from the predicted trajectory, and finally outputs the target location.

A. Normality Dection Module Based on Skewness and Kurtosis

In this paper, we propose a normality detection method based on skewness and kurtosis to divide the RSSI signal of the existing database to establish a more accurate WiFi fingerprint database. The main steps are shown in Fig. 2.

a) Normal distribution test: The analysis of skewness and kurtosis is used to determine whether the collected signal intensity values meet the normal distribution.
b) Normal model processing: If the normal distribution model is satisfied, we use the normal distribution function to estimate the probability density of the sample.
c) Kernel density function model processing: If the normal distribution model not satisfied, we use the kernel function to calculate the overall probability.
d) The mean value of the high probability signal value is stored in the WiFi fingerprint database.

Skewness and kurtosis are statistics that characterize the steepness and symmetry of the data distribution. By measuring the skewness coefficient, one can determine the degree and direction of data distribution asymmetry. Specifically, the skewness of random variable refers to the normalized third order standard central moment:

$$\gamma_1 = E\left[\left(\frac{E-\mu}{\sigma}\right)^3\right] = \frac{E[(x-\mu)^3]}{(E[(x-\mu)]^2)^{\frac{3}{2}}} \tag{1}$$

Kurtosis is:

$$\gamma_2 = E\left[\left(\frac{E-\mu}{\sigma}\right)^4\right] = \frac{E[(x-\mu)^4]}{\sigma^4} \tag{2}$$

Where X is a random variable, μ is the expected value of the random variable, and μ is the variance of the random

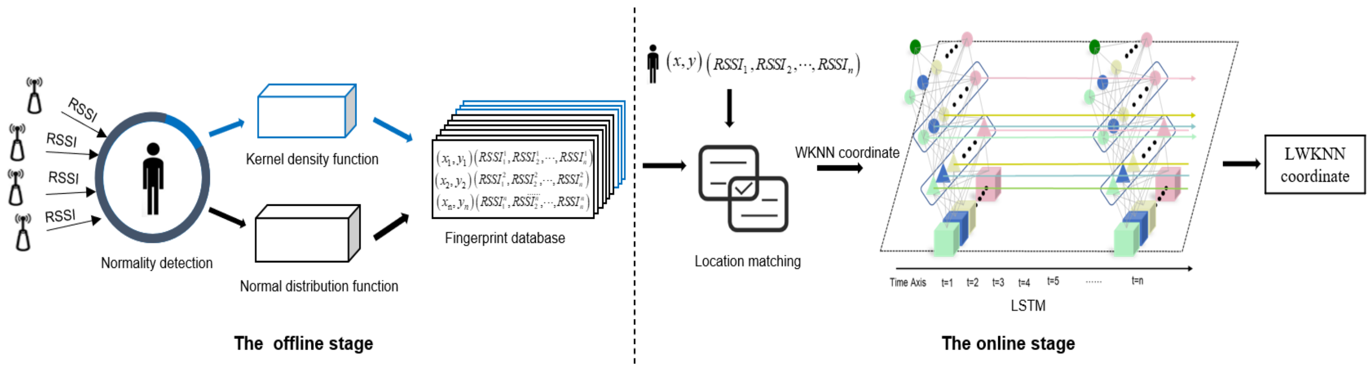


Fig. 1. Overview of the proposed framework.

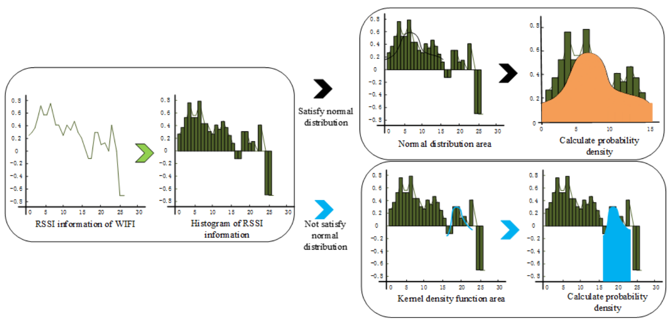


Fig. 2. Simulation diagram of signal acquisition and processing in offline stage.

variable. For example, if a random variable X follow the normal distribution, skewness $\gamma_1 = 0$ and kurtosis $\gamma_2 = 3$.

By testing whether the data samples obey the normal distribution, the sample data can be divided into two parts. The first part is the sample data obeying normal distribution, and the second part is the sample data not satisfying normal distribution model. Assume $X1, X2, , Xn$ represents the WiFi sample, $\bar{X}$ represents the average value of the sample, and $m_i = \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X})^i$ represents the sequential center distance of the sample. The skewness and kurtosis of the normal distribution are both 0. The equations for obtaining skewness and kurtosis are as follows:

$$b_s = \frac{E(X-EX)^3}{[Var(X)^{\frac{3}{2}})]} \tag{3}$$

$$k = \frac{E(X-EX)^4}{[Var(X)^2]} - 3 \tag{4}$$

The condition for using skewness and kurtosis normality tests is that all samples must have prior information that deviates from the normal state of skewness and kurtosis. Therefore, we assume that H_o: obeys the normal distribution, H_1: does not obey the normal distribution. Where H_o is the original hypothesis, the conditions are $b_s = 0$ and $k = 0$. The conditions of H_1 are: $b_s \neq 0$, $k \neq 0$. First calculate the statistics.

$$T = \frac{\hat{b}_s^2(n+1)(n+3)}{6(n+2)} + \frac{(\hat{k} + \frac{6}{(n+1)})^2(n+1)^2(n+3)(n+5)}{24n(n-2)(n-3)} \tag{5}$$

According to the statistical limit distribution, the degree of freedom is χ^2 distribution of 2, so its test rejection region is $\{T > \chi^2_{(1-\alpha)}(2)\}$, where $\chi^2_{(1-\alpha)}(2)$ is the distribution of $1-\alpha$ quintile degree of freedom of χ with 2.

As mentioned above, the RSSI samples collected in the offline phase are tested to determine whether the rejection domain of H_0 meets the requirements. If they are not satisfied, accept H_0 . When the significance level is $\alpha(0 < \alpha < 1)$, the overall sample is considered to obey the normal distribution. At this time, the probability density of the sample is about:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)}{(2\sigma^2)}\right] \tag{6}$$

Where $\mu = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[X_i - \mu]^2}$ are the parameters of the normal distribution, H_0 represents the RSSI value in the i^{th} sample, and n is the capacity of the RSSI sample.

If the total number of RSSI signal samples collected in the offline phase negates the H_0 hypothesis, it means that the overall distribution of the sample is significantly different from the normal distribution. Thus, the significance level is α, which means that the sample population does not obey the normal distribution, in which case we use the kernel density function to estimate the overall distribution of samples. Kernel density estimation is a method to estimate the overall probability density when the overall distribution is unknown.

Let $f(x)$ denote the probability density of the whole sample of X, and $X1, X2, , Xn$ denote the sample of population X, and there is a bounded function $k(y) \geq 0$ on the whole line. This function satisfies the following four conditions: $\int_{+\infty}^{-\infty} |k(y)|dy < +\infty$, $\lim\limits_{|n|\to+\infty} yk(y) = 0$, $k(-y) = k(y)$, $\int_{-\infty}^{+\infty} k(y)dy = 1$.

$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} k(\frac{x-X_i}{h})$ is the kernel estimate of the unknown probability density $f(x)$, where h is the kernel width and $k(y)$ is the kernel function.

We select kernel functions:

$$k(y) = \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{x^2}{2}\right] \tag{7}$$

The kernel function is used to estimate the probability density function for the entire sample

$$f(x) = \frac{1}{nh}\sum_{i=1}^{n}\exp\left[-\frac{(X_i - x)^2}{2h^2}\right] \tag{8}$$

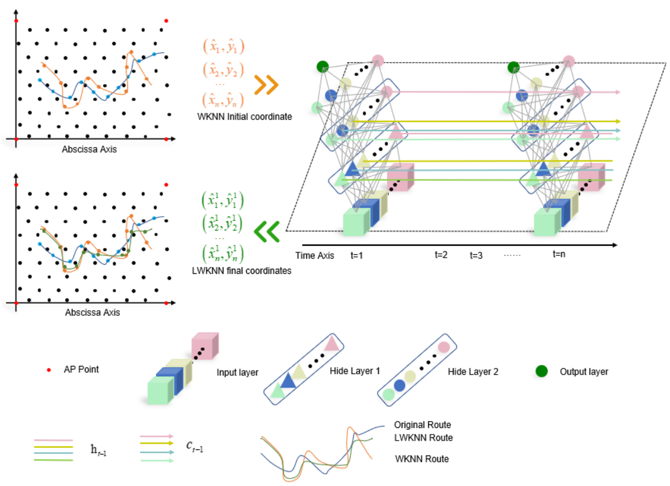


Fig. 3. Simulation diagram of signal acquisition and processing in offline stage.

The normality detection module based on skewness and kurtosis proposed above is for signal acquisition of the experimental environment in the offline phase. By screening out some sample data that does not meet the normal distribution, the kernel density function is used to estimate the overall distribution of the sample. Then, by using the Kalman filter to remove the impact of system errors, we can finally take the average as a determined RSSI value recorded in the fingerprint database.

B. LWKNN Prediction Positioning Module

In the online positioning stage, the traditional location fingerprint positioning algorithm based on RSSI mainly includes one of the following: NN, KNN, WKNN [18]. We propose the LWKNN algorithm, which is a fusion target trajectory prediction method based on the traditional WKNN. The WKNN method is used to locate the pedestrian's initial position by using the preprocessed fingerprint database. The initial coordinate position information is input into the input layer, and then we establish the LSTM network using this information. Finally, the final positioning result is obtained by training the positioning coordinates obtained by WKNN.

The LWKNN module proposed in this paper is shown in Fig. 3. This input layer input dimension is horizontal and vertical coordinates, that is, the dimension is 2. The hidden layer information h_{t-1} and c_{t-1} default to 0 at the initial time.

First, the mobile terminal collects the RSSI value of each AP point in real time, denoted as $x = [x_1, x_2, obtion.., x_n]$. Suppose the data in the fingerprint database is $X_i = [RSSI_1^i, RSSI_2^i, RSSI_3^i, ..., RSSI_n^i]$, where n represents the number of AP points ; $i \in [1, N]$, N is the number of records in the fingerprint database ; $RSSI_n^i$ denotes the signal strength of the n^{th} AP point collected by the i^{th} fingerprint point. Find the Euclidean distance between x and X_i :

$$L_i = \sqrt{\sum_{j=1}^{n}(x_i - RSSI_i^j) \times (x_j - RSSI_i^j)} \tag{9}$$

Marked as $L = \{L_1, L_2, ..., L_n\}$, sort the Euclidean distance from small to large, find the first $K(K \geq 2)$ fingerprint reference points in the sequence L , multiply the coordinates corresponding to each fingerprint reference point by a weighting factor and sum it up. The weighting factor is based on the matching value in the NN algorithm. The specific function is:

$$(\hat{x}, \hat{y}) = \sum_{i=1}^{K} \frac{\frac{1}{L_i(1+\alpha)}}{\sum_{j=1}^{K} \frac{1}{L_j(1+\alpha)}}(x_i, y_j), (k > 2) \tag{10}$$

Among them, $(\hat{x}, \hat{y})$ represents the final positioning result, L_n is the Euclidean distance between the node to be determined and the n-th fingerprint point, (x_i, y_j) represents the coordinates of the i^{th} reference point, and K represents the number of fingerprint reference points that are most matched. To prevent Eq. 10 from being meaningless, the value of α is as small as possible but not 0.

Next, we create a long short term memory network is established. A LSTM unit is divided into three gates, namely, forgetting gate, input gate and output gate. The network calculation is performed according to the following steps:

1) Determines whether part of the information is retained from the cell state, selectively forgetting the information from the previous cell state.
2) Decide to store information in the cell state and selectively record new information into the cell state.
3) Process the input of the current sequence position, determine the information that needs to be updated, and update the old cell state.
4) Determine the output content based on the content of cell state preservation, that is, selectively output the content of cell state preservation.

The result $(\hat{x}, \hat{y})$ of the initial WKNN positioning is passed as input to the LSTM network model, denoted as $\begin{bmatrix}\hat{x}_t \\ \hat{y}_t\end{bmatrix}$.

Its forgotten door is:

$$f_t = \sigma\left(w_f \begin{bmatrix} h_{t-1} & h_{t-1} \\ \hat{x}_t & \hat{y}_t \end{bmatrix} + b_f\right) \tag{11}$$

The input door is:

$$i_t = \sigma\left(w_i \begin{bmatrix} h_{t-1} & h_{t-1} \\ \hat{x}_t & \hat{y}_t \end{bmatrix} + b_i\right) \tag{12}$$

The input unit state is:

$$\wp_t = tanh\left(w_c \begin{bmatrix} h_{t-1} & h_{t-1} \\ \hat{x}_t & \hat{y}_t \end{bmatrix} + b_c\right) \tag{13}$$

$$c_t = f_t o c_{t-1} + i_t o \wp_t \tag{14}$$

The output door is:

$$o_t = \sigma\left(w_o \begin{bmatrix} h_{t-1} & h_{t-1} \\ \hat{x}_t & \hat{y}_t \end{bmatrix} + b_0\right) \tag{15}$$

Where σ is the sigmoid function, ω is the corresponding weight matrix, and b is the bias term.

The final output of LWKNN is:

$$\begin{bmatrix}\hat{x}_t^1 \\ \hat{y}_x^1\end{bmatrix} = o_t o tanh(c_t) \tag{16}$$

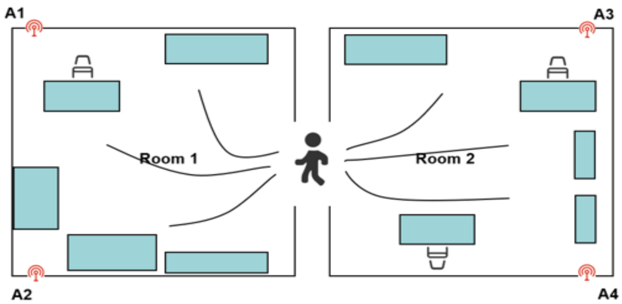


Fig. 4. Indoor data collection.

The LWKNN module proposed above integrates the LSTM model based on the WKNN, and finally combines the predicted trajectory coordinate information to obtain the pedestrian's final target location.

IV. Data and Experiments

A. Datasets

The datasets used in the experiment were published by Barsocchi et al. [23]. The overall floor plan of the acquisition is shown in Fig.4. During the data collection process, the smartphone remains at the chest level with the screen facing up. Each time the user is in a predefined location, the device will record the following additional data about the detected WIFI access point. The database includes: wireless network name, AP MAC address, AP received signal strength, where WiFi signal strength is expressed in dB.

B. RSSI Data Preprocessing Experiment

In IoMT, most medical sites are indoor with complex environments and a large number of medical personnel. WiFi received signal strength belongs to link layer information, which is susceptible to multipath propagation, human absorption, shadow effect and other factors in highly crowded medical settings. This paper analyzes the overall distribution of RSSI samples in indoor WiFi environment. The analysis of the RSSI sample data from the randomly selected dataset yields that a large proportion of the selected RSSI sample data have sample values that do not obey a normal distribution.

We use the above data set to verify the normality detection method based on skewness and kurtosis. By arbitrarily selecting 30 groups of RSSI samples from the target AP (receiving) point in the database and a single AP transmitting node, the sample data is preprocessed by the fusion of skewness kurtosis normality detection and Kalman filter. Its comparison with the mean model, median model, and normal distribution model is shown in Fig. 5.

The experimental results show that the skewness and kurtosis based normality detection method proposed in this paper can restore the ideal values well and eliminate the error values better in the offline stage. The high precision WiFi fingerprint database established on this basis can effectively improve the localization accuracy of the patients in medical sites during the localization phase.

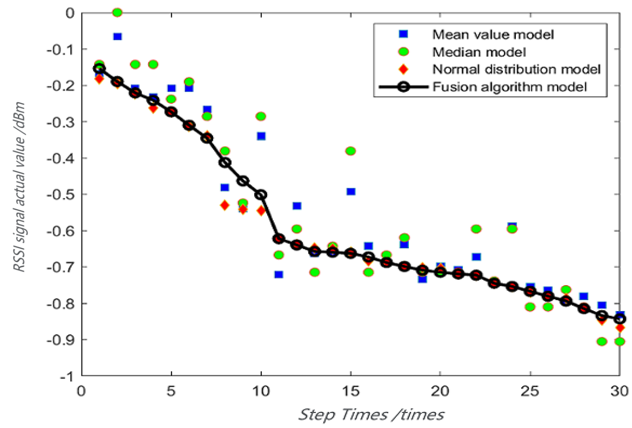


Fig. 5. Comparison of RSSI results of different models.

TABLE I
Average Error of Different Values of K When Using Normal Distribution Model

Value of K	Average Error $/m$
K=2	1.1483
K=3	1.0465
K=4	1.0317
K=5	1.1686

C. Experiment of Localization

In the fingerprint positioning algorithm WKNN, the positioning accuracy is the highest by comparing several experiments. Tab. I shows the average positioning errors for different values of K when using the normal distribution model. Therefore, we default the K-weighted proximity algorithm with a value of 4.

The trajectory prediction method proposed in this paper relies upon historical positioning information to establish the LSTM information base, so we obtained 100 sets of continuous positioning coordinate information through the WKNN positioning method. We verify the positioning accuracy of the LWKNN positioning algorithm using the LSTM information library. In this experiment, the first 80 coordinates in the continuous coordinate information are selected as the training set, and the remaining 20 coordinate information are used as input to compare the positioning effect of the two methods. As shown in Fig. 6, the positioning effect of LWKNN is better than that of WKNN and is closer to the true value. One of the reasons for the large error of WKNN is that the received unstable signals cannot be handled well during the online phase.

This paper uses the updated fingerprint database to compare the actual positioning error between the LWKNN method and the traditional WKNN method. We randomly select continuous LWKNN and WKNN coordinate information and pedestrian real coordinate information. We then import the coordinate information database, then calculate the difference between the LWKNN output and the real value, as well as the dif-

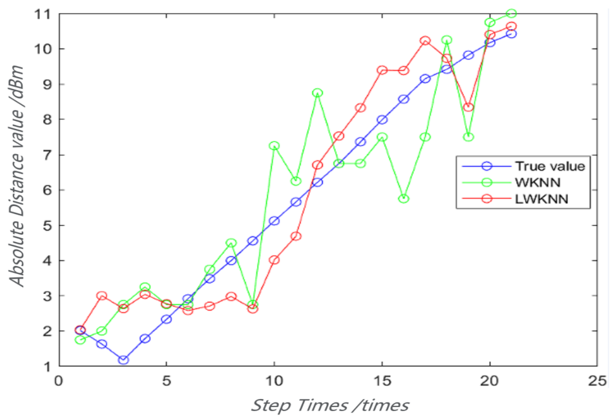


Fig. 6. Comparison of WKNN/LWKNN positioning effects.

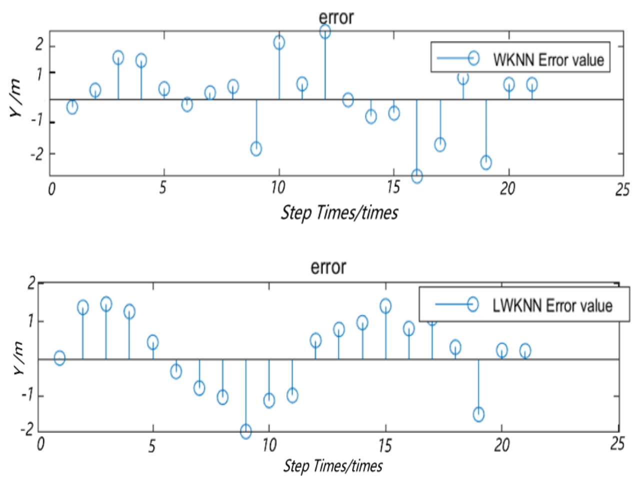


Fig. 7. Comparison result of positioning error value.

TABLE II
COMPARISON OF WKNN/LWKNN POSITIONING ERRORS

	WKNN	LWKNN
Maximum error/m	2.824	1.928
Average error/m	1.0317	0.8776
Maximum deviation/m	1.192	0.938
Standard deviation/m	1.3422	0.9961
Average deviation/m	1.3753	1.021

ference between the WKNN output and the real value, using MATLAB. The experimental results are shown in Fig. 7.

It can be seen from the above figure that the LWKNN method proposed in this paper can reflect the actual pedestrian real motion trajectory more smoothly than the traditional WKNN method, and has lower error. Compared with WKNN algorithm, the method proposed in this paper can effectively reduce the error and improve the localization accuracy in medical place environment.

By comparing the difference data between WKNN, LWKNN and real data, the following tables can be obtained.

From tab. II, it can be seen that the proposed localization algorithm based on fusing skewness and kurtosis of normality detection with LWKNN significantly outperforms the WKNN method just based on normality detection. The maximum error is reduced by 31.72%, the average error is reduced by 14.9%, the average deviation is reduced by 25.7%, and the maximum deviation is reduced by 21.3%.

V. CONCLUSION

The complexity of the current indoor environment in major hospitals makes it impossible to correctly estimate the distribution of RSSI samples measured by a single function. We propose two key steps to effectively improve positioning accuracy. First, we estimate the distribution of RSSI samples accurately in the offline phase and establish a more stable and accurate fingerprint database. The second part is to improve the positioning accuracy through efficient positioning methods in the positioning stage. In this paper, our proposed method, which fuses skewness and the kurtosis normality test with an improved WKNN method, effectively eliminates both the singular values in the collected RSSI signals and fluctuations in the data to establish a stable and accurate fingerprint database. The LWKNN algorithm effectively reduces the error impact caused by WiFi time variation and makes full use of historical information of pedestrian trajectory. Experiments show that our proposed method has lower error than both the traditional mean model and normal distribution model. Compared with the traditional WKNN method in IoMT, LWKNN can effectively improve the positioning accuracy of medical staff to patients and reduce potential safety hazards.

In the future, we will conduct research from the following two aspects. First, more deep learning techniques, especially convolutional neural networks (CNNs) in signal feature extraction, can be introduced to achieve finer feature extraction and create high-precision fingerprint databases. For the long short-term memory network model, we can also improve it to better incorporate historical trajectory information, so as to improve the accuracy of position prediction. These improvements will help to further improve the performance and application range of the LWKNN method in indoor positioning.

VI. ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China (62176067); Scientific and Technological Planning Project of Guangzhou (202103000040, 2023B03J1378); Research project of Guangdong Polytechnic Normal University (22GPNUZDJS14).

REFERENCES

[1] S. Jain, "Internet of medical things (iomt)-integrated biosensors for point-of-care testing of infectious diseases," *Biosensors and Bioelectronics*, vol. 179, p. 113074, 2021.

[2] A. Ghose, "Ubiheld: ubiquitous healthcare monitoring system for elderly and chronic patients," in *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pp. 1255–1264, 2013.

[3] L. R. M. Joyia G J and F. A, "Internet of medical things (iomt): Applications, benefits and future challenges in healthcare domain," *Journal of Communications*, vol. 12, pp. 240–247, 2017.

[4] C. Q. Liu Y and L. G, "ecosense: a hardware approach to on-demand sensing in internet of things," *IEEE Communications Magazine*, vol. 54, pp. 37–43, 2016.

[5] F. K. Wu Z and J. E, “A fast and resource efficient method for indoor positioning using received signal strength,” *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 9747–9758, 2016.

[6] L. T. N. Fang S H, “A dynamic system approach for radio location fingerprinting in wireless local area networks,” *IEEE Transactions on Communications*, vol. 58, pp. 1020–1025, 2010.

[7] L. J. Pei L and C. Y, “Evaluation of fingerprinting-based wifi indoor localization coexisted with bluetooth,” *Journal of Global Positioning Systems*, vol. 15, p. 3, 2017.

[8] B. A. Mazuelas S and L. R. M, “Robust indoor positioning provided by real-time rssi values in unmodified wlan networks,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, pp. 821–83, 2009.

[9] L. X. Y. Wang L J and P. W. J, “Wifi fingerprint location algorithm in indoor location based on signal mean filter+k-means+wknn,” *Microelectronics & Computer*, vol. 34, pp. 30–34, 2017.

[10] W. M. Cui X and L. J, “Indoor wi-fi positioning algorithm based on location fingerprint,” *Mobile Networks and Applications*, vol. 26, pp. 146–155, 2021.

[11] S. C. L. Chen K and C. J. B, “An indoor location fingerprint algorithm based on improved wknn,” *Navigation Positioning and Timing*, vol. 3, p. 58, 2016.

[12] N. S. Bai J, “Tests for skewness, kurtosis, and normality for time series data,” *Journal of Business & Economic Statistics*, vol. 23, pp. 49–60, 2005.

[13] M. Luo, “"wifi-based indoor localization using clustering and fusion fingerprint",” *2021 40th Chinese Control Conference (CCC). IEEE*, pp. 3480–3485, 2021.

[14] P. J. Wang J, “An enhanced indoor positioning algorithm based on fingerprint using fine-grained csi and rssi measurements of ieee 802.11n wlan,” *Sensors*, vol. 21, p. 2769, 2021.

[15] E. N. R. Guo X and A. N, “Robust wifi localization by fusing derivative fingerprints of rss and multiple classifiers,” *IEEE Transactions on Industrial Informatics*, vol. 16, pp. 3177–3186, 2020.

[16] X. Z. Huang B and J. B, “An online radio map update scheme for wifi fingerprint-based localization,” *IEEE Internet of Things Journal*, vol. 6, pp. 6909–6918, 2019.

[17] W. H. Xinchun L I, “Fingerprint location using sparse fingerprint acquisition and improved wknn,” *Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition)*, vol. 31, pp. 451–458, 2019.

[18] Y. L. Kong G and S. D, “Research on an algorithm of fingerprint location based on k-means and wknn,” *GNSS World of China*, vol. 41, pp. 89–92, 2016.

[19] H. S. Pan S and P. D, “Wireless localization method based on ahp-wknn and amendatory akf,” *Wireless Communications and Mobile Computing*, pp. 1–11, 2021.

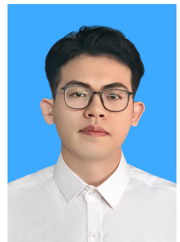
[20] L. Z. Zhao Z and W. R, “I-wknn: Fast-speed and high-accuracy wifi positioning for intelligent stadiums,” *Computers & Electrical Engineering*, vol. 98, 2022.

[21] D. Z. Li X and H. L, “A k-nearest neighbor indoor fingerprint location method based on coarse positioning circular domain and the highest sim-ilarity threshold,” *IOP Publishing Ltd*, 2022.

[22] C. G. Cao X, “An improved fingerprint indoor localization algorithm based on combination weight,” *Bulletin of Surveying and Mapping*, vol. 2, p. 6, 2018.

[23] S. C. C. G. D. Bacciu, P. Barsocchi and A. Micheli, “An experimental characterization of reservoir computing in ambient assisted living applications,” *Neural Computing and Applications, Springer-Verlag*, vol. 24, pp. 1451–1464, 2014.

Xu Lu is a Professor in the School of Computer Science, Guangdong Polytechnic Normal University, China. He received the B.S. degrees from Nanchang University, Jiangxi, China, in 2006, and the M.E. and Ph.D. degree from the Guangdong University of Technology, Guangdong, China, in 2009 and 2015, respectively. His research interests include image processing, artificial intelligence and smart system.



Kejie Zhong is currently pursuing a master’s degree in Systems Engineering at the School of Computer Science, Guangdong Polytechnic Normal University. His main research directions include the Internet of Things and artificial intelligence.

Zhiwei Guan is currently pursuing a master’s degree in Systems Engineering at the School of Computer Science, Guangdong Polytechnic Normal University. His main research directions include the Internet of Things and artificial intelligence.

Jun Liu received his M.S. and Ph.D. degrees in Control Theory and Control Engineering from Guangdong University of Technology, China in 2012 and 2015, respectively. He is currently an associate professor in the School of Automation, Guangdong Polytechnic Normal University, China. He is working in the Internet of things and wireless sensor networks. E-mail:liujun7700@163.com