Received May 15, 2016, accepted May 19, 2016, date of publication June 15, 2016, date of current version July 7, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2579879

Fingerprint and Assistant Nodes Based Wi-Fi Localization in Complex Indoor Environment

QIYUE LI[1], (Member, IEEE), WEI LI[1], WEI SUN[1], JIE LI[2], AND ZHI LIU[3], (Member, IEEE)

[1]School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 23009, China
[2]School of Computer and Information, Hefei University of Technology, Hefei 23009, China
[3]Global Information and Telecommunication Institute, Waseda University, Tokyo 169-8555, Japan

Corresponding author: J. Li (lijie@hfut.edu.cn)

This work was supported in part by the Anhui Province Key Laboratory of Affective Computing and Advanced Intelligent Machines under Grant ACAIM201505 and in part by the National Natural Science Foundation of China under Grant 61301114 and Grant 51304058.

ABSTRACT With the extensive development of Wi-Fi, indoor location services based on received signal strength (RSS) fingerprints have attracted increasing attention from researchers. In complex indoor environments, multipath and non-line-of-sight (NLOS) conditions would lead to large errors in measured values, thereby reducing indoor positioning accuracy. In this paper, we propose a Wi-Fi indoor localization method based on collaboration of fingerprint and assistant nodes. First, appropriate assistant nodes based on the similarity of RSS sequences are elaborately selected around the unknown node and distances between them are used as auxiliary information to improve the positioning accuracy. Furthermore, in the complex indoor circumstances that result in NLOS error, an adaptive Kalman filter with colored noise is used to mitigate the time-of-flight ranging error. Experiments demonstrate that in complex indoor environments, our system can outperform its counterparts with robust performance and low localization estimation error.

INDEX TERMS Indoor localization, searching model, TOF, Wi-Fi.

I. INTRODUCTION

With the booming development of mobile Internet and intelligent terminals, due to its imponderable social and commercial values, localization-based services (LBSs) have attracted considerable attention from researchers. The demand for an indoor positioning service or indoor LBS (iLBS) has also accelerated given that people spend the majority of their time indoors. However, the complexity of the indoor environment often severely affects the accuracy of indoor localization. Complications include non-line-of-sight (NLOS) reference objects, the presence of obstacles, signal fluctuation or noise and environmental change [1]–[4].

The indoor localization problem has been met with much research and study. In particular, the ubiquity of wireless fidelity (Wi-Fi) networks has made it a target of localization efforts especially due to its pervasiveness in infrastructure and mobile clients [5]. Many indoor localization technologies have been promoted such as RFID, Wi-Fi and Bluetooth. Nevertheless, due to the limitations of the hardware of a common smart cell phone and the facilities that are installed in open environments, location services based on Wi-Fi are widely used in indoor navigation applications [6], [7].

Wi-Fi based localization research typically falls into two distinct paradigms: radio propagation modeling and radio frequency (RF) fingerprinting [8]. The former attempts to infer the position of the target device using equations that express wireless signals in the specific environment and obviously it is very important and difficult of wireless signal propagation parameter estimation to obtain an accurate position due to the complex signal transmission pattern in indoor environment [9]. Research like [9] has indicated that the well-known log-normal shadowing model complies with the following formula (1).

$$\mathrm{PL}\,(d) = \overline{PL\,(d_0)} - 10nlog_{10}\left(\frac{d}{d_0}\right) + X_\sigma \qquad (1)$$

where d_0 means the near-Earth reference distance (which is determined experimentally) and d is the distance between the transmitter and receiver. $\overline{PL\,(d_0)}$ represents the received strength when the distance is d_0 and the same goes for $PL\,(d)$ when the distance is d. And the coefficient n indicates the rate of the loss exponent which varies with the complexity of environment and is always set to 2∼4 under indoor circumstances. X_σ is to model the shadowing effect, which follows a Gaussian distribution (in dB) with mean zero and standard deviation σ.

The latter has focused on Wi-Fi fingerprint-based positioning systems using RSS [10]–[13]. The fingerprinting positioning method results from pattern recognition. Theory suggests that with increasing propagation distance, factors such as indoor obstacle blocking and interference from other wireless devices in a complex environment will attenuate the transmission of a signal from an Access Point (AP) and will make signal strength exhibit non-regular distribution [10]. APs generally could be observed simultaneously by terminal devices in regions covered with WLAN, and the signal strength of each AP subject to the same characteristic [11]. Therefore, the terminal device at each specific localization can observe a group of APs and measure the RSS values of different APs, however RSS values received from APs set at different positions are different [12], [13]. So the RSS information from the APs described above is called a "fingerprint" [14].

Usually APs periodically broadcast a beacon signal that demonstrates the existence of APs. Therefore terminal devices equipped with a wireless network card can obtain three types of information about APs: APs' names, MAC addresses and RSS values through the signal transmitted from APs, even though they do not communicate with the WLAN. Using APs' names or MAC addresses, terminal devices could distinguish RSS values from different APs.

The RSS-based fingerprinting localization algorithm does not require the path-loss model of signal propagation. The positioning process seen in Fig. 1 typically falls into two stages.

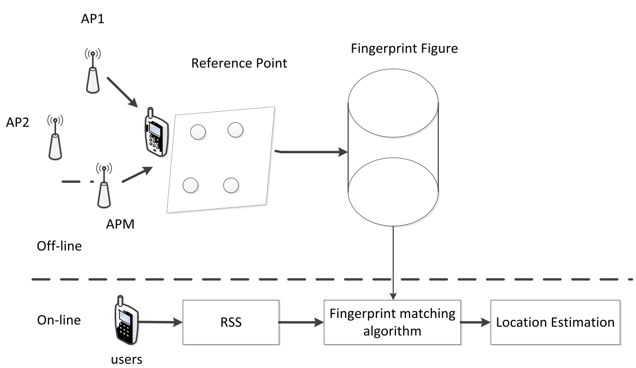


FIGURE 1. The scheme of fingerprint localization algorithm.

Off-line phase: a two-dimensional Cartesian coordinate system is established in the indoor environment, and reference points (RPs) are set at a fixed interval according to the desired positioning accuracy. Next, a RSS signature map (which is also called an RSS locating fingerprint database) containing the RSS values corresponded to RPs' coordinates is effectively built by traversing the area. In order to obtain the database mentioned above, we need to set some RPs in the positioned region and then an RSS data collection device is used to traverse all RPs while recording RSS values and coordinates.

On-line phase: for the sake of acquiring the unknown node's localization, its RSS values are measured, and then matching algorithms are utilized to compare it with the data stored in the database.

The position accuracy of the traditional fingerprint-based localization algorithm depends on the accuracy of the matching algorithm and fingerprint databases. However, in complex indoor environments impacted by signal reflection, refraction, and obstacles, RSS value has a greater error due to large fluctuation, which seriously affects the fingerprint database's accuracy. Nevertheless, traditional localization algorithms only utilize single node while neglecting other indoor nodes' information. In fact, current indoor environments (e.g., offices, shopping malls) are filled with Wi-Fi devices. Therefore, our intuition is that these Wi-Fi devices' information may be used to improve positional accuracy; they can be considered assistant nodes. Hence this paper proposes a novel Wi-Fi indoor localization algorithm based on the collaboration of RSS and assistant nodes.

We select assistant nodes based on the principle of RSS similarity, which will be introduced in the following section. After calculating the values of RSS similarity between the unknown node and surrounding ones, we can select assistant nodes effectively. Next, we measure distances between each pair of Wi-Fi devices through time-of-flight (TOF) range, which can be used as auxiliary information to improve positional accuracy. The TOF-based methods measure the distance based on estimation of the signal propagation delay, *i.e.*, the TOF, between a transmitter and a receiver since in free space or in air, radio signals travel at the constant speed-of-light. We adopt an adaptive Kalman filter to eliminate NLOS ranging error and build a searching model with an exhaustive and maximum gradient algorithm to obtain an accurate localization for the unknown node.

In light of the great demand for indoor localization services, we propose a novel Wi-Fi indoor localization algorithm based on the collaboration of RSS and assistant nodes in this paper. It not only utilizes the RSS fingerprint database but also selects assistant nodes close to the unknown node in space by comparing the RSS values' sequence. And then we take the distances between all the assistants and the unknown node into account. Aiming at the severe NLOS error caused by the complicated indoor environment, an adaptive Kalman filter with colored noise is used to eliminate the ranging error to enhance localization accuracy. Based on the additional ranging information of assistant nodes, a searching model is established to optimize localization error and reduce the effects of the environment's complexity.

Unlike the traditional fingerprint-based algorithm, this paper utilizes many other devices surrounding with the unknown one to improve positioning accuracy. Furthermore, NLOS mitigation method is applied in our algorithm to eliminate measurement error caused by complex indoor environment. That can give a better localization results than the current literatures which just simply extend the single device localization method into a multiple devices version, designating weights to each device according to matching result against the fingerprint database or distances measurements between the devices.

It is worth to note that there is no essential difference between node and device, and in fact we use the two notations interchangeably, while the unknown node means the interested device whose location is unknown and needs to be estimated.

The rest of this manuscript is organized as follows. Section II reviews the related work on Wi-Fi localization algorithms, and in Section III, the theoretical basis is explained, and the system structure is depicted in detail. In Section IV, experimental results are presented, and the paper's conclusions are presented in Section V.

II. RELATED WORK

Considerable research has been performed on indoor localization. Paolo Barsocchi proposes a novel localization algorithm which selects and weights the RSS measurements according to their strength, and uses an automatic training that only exploits information from the anchors, without requiring human input [15]. However, information from anchor nodes in this algorithm is not taken into account and the propagation model of the wireless signal is simplified which affects the localization accuracy when the system is deployed in a complex environment. Mussa Bshara proposes using fingerprinting localization depending on RSS-based observations for positioning and tracking in wireless networks with a new approach called the base-station-strict (BS-strict) methodology [16]. However, the time information (stamps) available with the measurements are not considered. Nevertheless, Ramsey Faragher examines the unique properties of BLE (Bluetooth Low Energy) signals and studies the application of fingerprinting to locate BLE devices in an environment with BLE beacons without considering the effect of people walking and the number of APs on the positioning accuracy [17].

Researchers have shown that matching algorithms and the RSS data's accuracy are the most important factors in fingerprint based positioning systems [18]. Concerning the problems mentioned above, Chung-Hao Huang proposes Kalman-filter drift removal and Heron-bilateration localization estimation to effectively reduce RSS drift and localization error, without any sacrifice of localization granularity and accuracy [19]. Jianwei Niu detects WiFi APs to form WiFi fingerprints from the signals collected by ZigBee interfaces, and aligns a pair of fingerprints with a matching algorithm to improve the localization accuracy [20], [21].

However, this work does not take the effect of NLOS on distance measurement into account.

Maxim Shchekotov describes Wi-Fi lateration methods based on signal propagation models and signal strength data collection for indoor localization [22]. The proposed methods use a log-normal path loss model for signal propagation and RSS measurement collection for distance estimation and a lateration approach for localization. Furthermore a traditional radio map is replaced by special ring radio map building. However, noise in the measured RSS is assumed to be a zero mean Gaussian random variable with variance RSS, and the approach is only suitable for one room if the number of APs and their density is low.

To addresses the implementation of indoor real-time localization systems (RLTS), Saverio Pagano presents a RSS-based tag based on WSNs with an anchor nodes clustering strategy based on the Euclidean distances to reduce the number of anchor nodes and to increase the beacon node localization accuracy [23]. Yet, the influence of unfixed distance between the beacon node and anchor nodes on localization accuracy has not been thought out. Unlike the traditional approaches, Chouchang Yang proposes a Wi-Fi based positioning technique utilizing the transmission of multiple predefined messages [24]. However, this work assumes that each AP has four antennas and the user device is a smartphone with a single antenna without considering the effects of the number of each AP's antennae on the localization.

Zhao Fang proposes a novel gradient-based AP localization approach, inferring the direction (minus gradient) of an AP just from the locally received signal strength variations [25]. This work introduces a direction clustering method to identify and eliminate the inaccurate gradients, but only selects partially accurate gradients to perform optimization. Jongtack Jung proposes a range-based RSS localization algorithm consisting of DB-assistance, a ration based algorithm, and an elementary machine learning algorithm [26]. However, the environmental influences on distances converted from RSS measurements is not considered without related analysis. Yuan Zhuang proposes a novel algorithm for automatic estimation of AP localizations, also for propagation parameters in dynamically changing indoor environments without requiring some knowledge of the propagation parameters in advance [27]. However, this work does not consider the influence of the variable caused by shadowing not submitting to Gaussian distribution.

The localization accuracy of the traditional fingerprint-based algorithm depends on the matching algorithm and fingerprint databases. However, in complex indoor environments impacted by signal reflection, refraction, and obstacles, RSS value has a greater error due to large fluctuation, which seriously affects the fingerprint database's accuracy. Furthermore, only the unknown node's information is used in the traditional localization algorithm while many surrounding devices are neglected which ban be utilized to improve the positioning accuracy.

Current localization algorithms which take the surrounding devices into account just simply extend the single device localization method into a multiple devices version, designating weights to each device according to matching result against the fingerprint database or distances measurements between the devices. However, they do not consider whether a certain surrounding device can be used to improve the positioning accuracy. Furthermore, under the complex indoor environment, the distance measurements may vary from time to time due to the changing wireless signal propagation and how to minimize the ranging error is also not revealed in the current literatures.

In this paper, we elaborately select appropriate assistant nodes around the unknown one based on the similarity of RSSI sequences and study how to use the auxiliary information to improve the positioning accuracy. Furthermore, we also present a method based on the adaptive Kalman filter with colored noise to mitigate the TOF ranging error in the complex indoor circumstances.

III. Wi-Fi LOCALIZATION SYSTEM BASED ON COLLABORATION OF FINGERPRINT AND ASSISTANT NODES

A. ASSISTANT NODES INTRODUCTION

With the rapid development of society, various electronic devices are used with the result that current indoor environments seriously affect the propagation of radio signals. A number of factors could attenuate and even block a signal, such as a node's feature, the antenna's orientation, or the environment around the node (e.g., walls, floors, obstructions).

Traditional RSS-based fingerprint localization algorithms only take advantage of the unknown node while neglecting other Wi-Fi devices. Around the unknown node, we can select several nodes to provide additional information to improve the positioning accuracy. The intuition underlying our design is to utilize the nodes close to the unknown one. Does this intuition make sense?

We did experiments to compare different devices' RSS values sequence in the environment as illustrated in Fig. 2 which reflects a complicated environment including many APs and Wi-Fi devices deployed in our lab. There are also many obstacles which affect signal propagation. In the experiments we measured the RSS data received by nodes 1, A1 and A4 from AP1, which is plotted in Fig. 3, and compared their respective fluctuations in RSS value under the same circumstances.

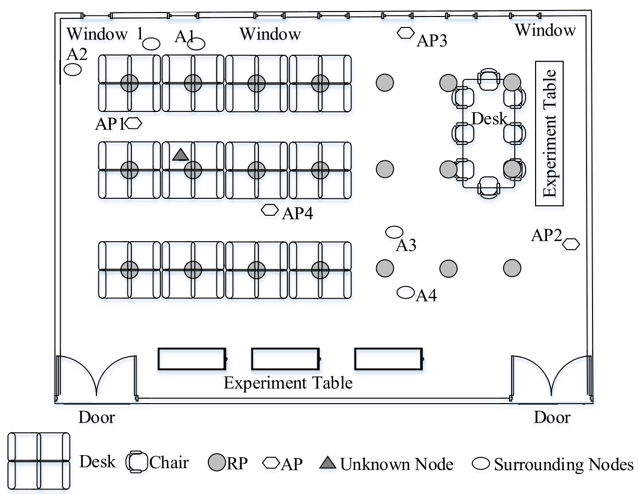


FIGURE 2. The scheme of nodes' spatial positions.

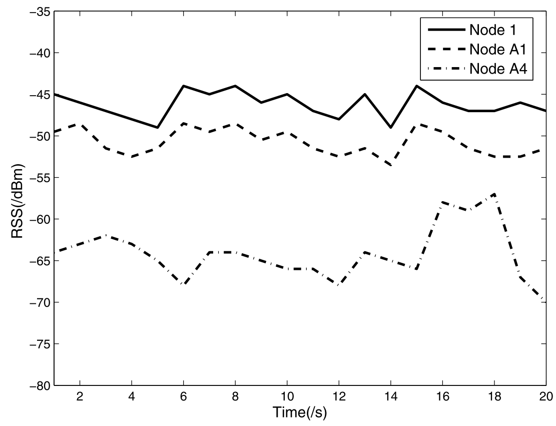


FIGURE 3. The RSS values of Nodes l, A1 and A4.

From Fig. 3, we can see that node l is close to A1 and A2, while far from A3 and A4 and that the RSS data of node 1 is very similar to that of A1 but totally different from that of A3. That's because the nearby nodes experience a similar multipath environment (e.g., reflectors and obstacles in the environment). Thus if the nodes surrounding the unknown one are selected as the assistant nodes to provide additional information in a fingerprint localization algorithm, the positioning accuracy can be improved due to the similar error repercussion. In addition, it has been proved that the Cramer-Rao lower bound (CRLB) of the unbiased estimation can be reduced when a closer AP is utilized rather than a distant one [28]. Based on the above reasons, we selected the nodes close to the unknown one as the assistant nodes and derived from the above experiments an effective process for selecting assistant nodes by comparing the similarity of RSS data sequences. This will be described in the following section.

B. SYSTEM OVERVIEW

Current indoor environments contain many Wi-Fi devices that can provide beneficial auxiliary information, so we select them as assistant nodes based on RSS similarity in indoor circumstances and then use them to acquire the unknown node's coordinates and improve localization accuracy. Then, distances measured between each pair of nodes are used as additional information to improve localization accuracy, and we apply TOF ranging to measure distance among Wi-Fi devices using the method mentioned like [29].

However, TOF range error is not subject to Gaussian distribution [30], in complex indoor environments. For the sake of eliminating NLOS error, an NLOS error mitigation algorithm based on an adaptive Kalman filter with colored measurement noise is presented. Therefore, we establish a model of colored noise and dynamically adjust the filter parameters based on the severity of the NLOS environment.

Meanwhile, a searching model is built according to the distances measured above and exhaustive and maximum gradient algorithms are applied to optimize localization estimation error. Combing these methods, this paper presents a novel

Wi-Fi indoor localization algorithm based on RSS and assistant nodes collaboration.

The overall proposed scheme is shown in Fig. 4.

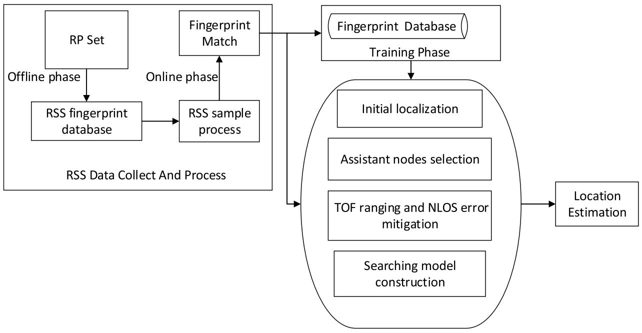


FIGURE 4. **The indoor localization model.**

Our system works as follows: first we build the fingerprint database on the off-line stage and then assistant nodes are acquired on the on-line stage. Next, we calculate the initial coordinates of the unknown node and assistant nodes, and measure distance among them, eliminating NLOS range error. Eventually, we build a searching model to acquire the final accurate coordinates of the unknown node.

1) INITIAL LOCALIZATION

Based on the RSS fingerprint database, the coarse-grained position of the unknown node and surrounding ones can be acquired with a traditional localization algorithm. Distances between on-line RSS sampling values and off-line RSS values stored in the fingerprint database are defined as formula (2).

$$D_i = \left(\sum_{j=1}^{n} \left|RSSI_j - RSSI_{ij}\right|^2\right)^{1/2} \quad i = 1, 2, 3 \ldots m \quad (2)$$

where the numbers of RPs and APs are m and n, respectively. $RSSI_{ij}$ represents j AP's RSS value stored in fingerprint database, and $RSSI_j$ also stands for the j AP's RSS value at on-line phase. Next, based on the KNN algorithm [31], k RPs meeting for acquirement are obtained and their respective coordinates are (x_i, y_i) $(i = 1, 2, 3 \ldots k)$, therefore we acquire the unknown node's initial localization through formula (3), which means the coarse-grained estimated localization of the unknown node is the centroid of the k nearest RPs.

$$(\overline{x}, \overline{y}) = \frac{1}{k}\sum_{i=1}^{k}(x_i, y_i) \quad (3)$$

2) ASSISTANT NODES SELECTION

As mentioned above, it is shown that when two nodes are deployed close in space, they will experience almost the same multipath environment and their RSS values' tendency is very similar. Thus we can select assistant nodes by comparing the similarity of RSS data sequences.

As shown in Fig. 5, node A is close to node B with n APs surrounding (represented by solid and dashed lines, respectively). In order to obtain their RSS similarity we

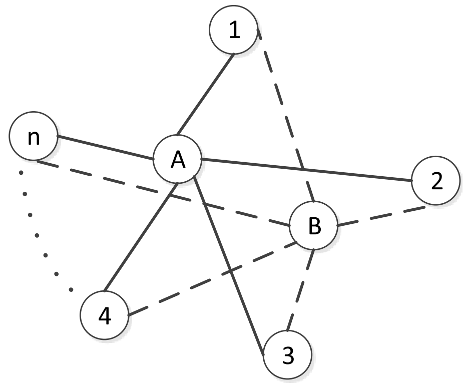


FIGURE 5. **The scheme of RSS similarity.**

collect their respective RSS values which is shown as $RSSI_A = (RSSI_{A1}, RSSI_{A2} \ldots RSSI_{An})$ and $RSSI_B = (RSSI_{B1}, RSSI_{B2} \ldots RSSI_{Bn})$.

In Fig. 3, it can be observed that in spite of the first two lines sharing similar trends, they are asynchronous in the time axis, which means that they reach their maximum and minimum values at different times. Thus it is unsuitable to use Euclidean distance to measure the similarity of the two time series of RSS data. In order to solve this problem, we are inspired by the word matching problem in speech recognition [32].

Even if the same person says the same word twice, she may intonate and speed up differently causing the same word to be stretched and scaled differently across clips. The speech recognition community matches such warped time signals using a technique called Dynamic Time Warping (DTW), which we adopt for our problem. We use DTW instead of Euclidean distance to measure the similarity of the two time series of RSS data.

The basic principles of DTW are explained as follows: in order to distinguish the similarity between two sequences, we need to bend the two sequences at local time which is beneficial for searching the optimal alignment of the two sequences. Compared with Euclidean distance, DTW distance (mean optimized path) effectively takes advantage of time bending and could reduce the influence due to different sampling time, which could better reflect the logic similarity of two sequences. If two labels are close in space, the impact of radio due to surroundings is the same. And if the RSS similarity of two labels is large, we likely consider them close to each other. Therefore DTW distance can be applied to distinguish the similarity of two labels close in space, which helps us determine the spatial position of the two labels. For the sake of eliminating the influence resulting from different sampling time, DTW distance between two nodes is defined as formula (4).

$$\begin{aligned} D_{dtw} &= D_{base}\left(RSSI_{A1}, RSSI_{B1}\right) \\ &\quad + min\left\{\begin{array}{l} D_{dtw}\left(RSSI_{A1}, RSSI_{B1}\left[2:-\right]\right) \\ D_{dtw}\left(RSSI_{A1}\left[2:-\right], RSSI_{B1}\right) \\ D_{dtw}\left(RSSI_{A1}\left[2:-\right], RSSI_{B1}\left[2:-\right]\right) \end{array}\right\} \end{aligned} \quad (4)$$

In this paper, Minkowski distance is used as base distance (D_{base}) and with D_{dtw} decreasing, distance between two nodes is closer which shows that node A is closer to node B in space. If $D_{dtw} > \delta$, the surrounding node can be selected as an assistant node of the unknown one and δ is usually recognized as empirical value.

3) TOF RANGING AND NLOS MITIGATION

In this paper, distances between the unknown node and assistant nodes are measured as auxiliary information to improve positioning accuracy. For this study, we use the same TOF ranging method [29]. However, the TOF measurements change significantly when some objects shadow the wireless link in complex indoor NLOS environments [28]. Obviously, this NLOS error changes with the environment, and its distribution is often difficult to obtain. To eliminate the NLOS error, a Kalman filter is frequently used which originally is designed to smooth White Gaussian noise. However, it is proven that TOF ranging error caused by multipath and NLOS propagation is not subject to Gaussian distribution [29]. Due to this mismatch, the traditional Kalman Filter is not sufficient enough to limit the huge ranging error.

Thus, we present an algorithm for NLOS error mitigation based on an adaptive Kalman filter with colored measurement noise to eliminate NLOS error. A colored noise model is firstly established according to measurement noise and the filter parameters are adjusted dynamically based on the severity of the NLOS environment.

a: THE PROCESS OF KALMAN FILTER WITH COLORED NOISE

Considering the linear system with state space representation, the system state equation and observation equation is defined as follows:

$$x_{k+1} = Ax_k + B\alpha_k \tag{5}$$

$$r_k = Cx_k + \beta_k \tag{6}$$

where A, B, and C are known constant matrices, α_k is process noise and β_k is ranging error. In (6) r_k is a one-dimensional observation vector and x_k is defined as:

$$x_k = \begin{bmatrix} d_k & \dot{d}_k \end{bmatrix}^T \tag{7}$$

where d_k denotes distance to be estimated and $\dot{d}_k$ denotes the first derivative of d_k and β_k is colored noise. Colored noise of the k times consists of the previous noise and zero-mean White Gaussian Noise [33]. Therefore, β_k can be given by

$$\beta_k = N_{k-1}\beta_{k-1} + \gamma_k \tag{8}$$

where N_k is an auto-regression coefficient and γ_k is White Gaussian Noise.

Thus the entire algorithm of the Kalman filter with colored noise can be described as follows:

$$P_{k,k-1} = AP_{k,k-1}A^T + BQB^T \tag{9}$$

$$H_{k-1} = [CA - N_{k-1}C] \tag{10}$$

$$G_k = \left(AP_{k-1,k-1}H_{k-1}^T + BQB^T\right) \cdot \left(H_{k-1}P_{k-1,k-1}H_{k-1}^T + CBQB^TC^T + \hat{R}_{k-1}\right)^{-1} \tag{11}$$

$$P_{k,k} = (A - G_kH_{k-1}) \cdot P_{k-1,k-1}A^T + (I - G_kC)BQB^T \tag{12}$$

$$\hat{x}_{k/k} = A\hat{x}_{k-1/k-1} + G_k \cdot (v_k - N_{k-1}v_{k-1} - H_{k-1}\hat{x}_{k-1/k-1}) \tag{13}$$

where $P_{k,k-1}$ and $P_{k,k}$ denote the predicted error covariance matrix and the estimated error covariance matrix, respectively, H_k denotes the coefficient matrix, Q denotes the noise covariance matrix, $\hat{R}_k$ denotes the ranging error covariance matrix, G_k and $\hat{x}_{k/k}$ denote the filtering gain and estimated state of the k times. In this algorithm, the observation vector is obtained by TOF ranging and the TOF distance measurement error is estimated first. The state is estimated through the algorithm in the end. From the entire process, it can be observed that $\hat{x}_{k/k}$ relates with the current input and previous estimated state, so real time processing can be realized only with the previous saved state value.

b: KALMAN FILTER WITH ADAPTIVE COLORED NOISE

We have presented the iterative process of the Kalman filter with colored noise in the above, in which the auto-regression coefficient N_k is known in advance. However, in realistic environments, movement of nodes, changes of surrounding environments, and complex wireless propagation paths all result in ranging error difficult to predict and model. It is clear that the N_k of ranging error is difficult to obtain in realistic environments. In the following, we will show how to estimate N_k and combine the estimation with the Kalman filter with colored noise. In addition, the change of ranging noise covariance R_k reflects the movement of ranging error, so we can utilize R_k to compute N_k. According to the Kalman filter and mathematics, the formulas can be given by

$$\hat{x}_{k/k-1} = A\hat{x}_{k-1/k-1} \tag{14}$$

$$z_k = r_k - C\hat{x}_{k-1/k-1} \tag{15}$$

$$\bar{z}_i = \frac{1}{i}\sum_{j=1}^{i} z_j \tag{16}$$

$$\hat{S}_k = \frac{1}{k-1}\sum_{i=1}^{k}(z_i - \bar{z}_i)(z_i - \bar{z}_i)^T \tag{17}$$

$$\hat{R}_k = \hat{S}_k - CP_{k,k-1}C^T \tag{18}$$

where $x_{k/k-1}$ and z_k are defined as the $\boldsymbol{k}$th predicted distance and innovation value, respectively. γ and β are independent of each other, and according to their covariance we can find

$$\hat{R}_k = N_{k-1}^2\hat{R}_{k-1} + var(\gamma_k) \tag{19}$$

where $var(\gamma_k)$ is the covariance of γ_k and $\hat{R}_k$ is the estimation of R_k. For the sake of simplicity, it is generally assumed that $var(\gamma_k)$ is equal to Q and NLOS error is positive and much larger than noise, and therefore N_k is also positive. In the actual process of filtering, it is possible that $\hat{R}_k$ is less than $var(\gamma_k)$. It is obvious that the possible outcome does not match the algorithm. Therefore, we set $N_{k-1} = 0$ in that case, which does not affect the result of the filter. Based on the above, we have

$$N_{k-1} = \left\{ \begin{matrix} \sqrt{\dfrac{\hat{R}_k - var(\gamma_k)}{\hat{R}_{k-1}}} & \hat{R}_k > var(\gamma_k) \\ 0 \quad otherwise & \end{matrix} \right\} \tag{20}$$

Thus, we can conclude that the whole process of the Kalman filter with adaptive colored noise is based on formulas (8) $\sim$ (13) and the auto-regressive coefficient N which can be computed by formulas (14) $\sim$ (20). In order to show the whole process better, we draw the overall flow diagram of the Kalman filter with adaptive colored noise shown in Fig. 6.

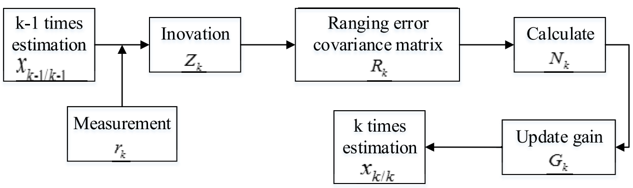


FIGURE 6. The overall diagram of the Kalman filter with adaptive colored noise.

4) MODELING AND SEARCHING

Now we can seek the optimal position of the unknown node. After the initial localization process described in Section III Subsection B using KNN algorithm [31] is conducted, and a coarse-grained initial guess of the unknown nodes' and assistant nodes' positions could be obtained. After the NLOS error mitigation algorithm is performed, the TOF-based distance measurements are accurate, which make the spatial model, including the unknown node, and all of the assistant nodes intact. Then the fixed structure with additional ranging information can be used to estimate the optimal position based on coordinates obtained in the initial localization process. Under these circumstances, we can search for the ultimate localization result by setting a searching scope (circle of radius R) around the initial position. During searching process, the spatial structure consisted of the unknown and assistant nodes keeps fixed and can be moved in the scope, as illustrated in Fig. 7.

The ultimate localization result can be obtained from the following objective.

$$\operatorname{argmin} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} (RSS_{ij}^{M} - RSS_{ij}^{F})^2 \right) \tag{21}$$

where m means the total number of the unknown and assistant nodes selected, and n means the number of deployed APs. RSS_{ij}^{M} is RSS data measured by each node, while RSS_{ij}^{F} is RSS data calculated against the fingerprint database with

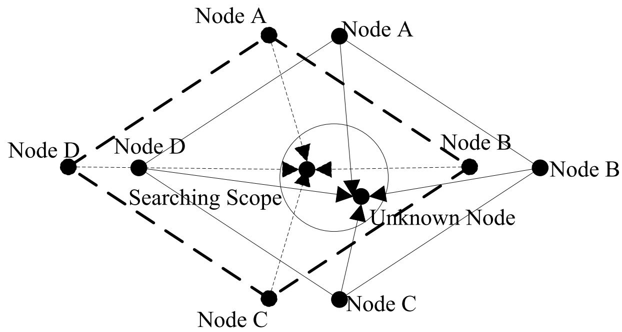


FIGURE 7. Fixed structure with four assistant nodes.

each node's coordinates using KNN algorithm similar to [31] during the searching process. Equation (21) produces the final results by finding locations of each node against the fingerprint database, by minimizing the sum of RSS distances, each of which is between RSS measurement data and the signature data at different location in the searching scope. In the interest of acquiring the ultimate localization, two solving methods are applied as follows in this paper.

a: EXHAUSTIVE ALGORITHM

In exhaustive searching method, the circle mentioned above would be divided into α equal parts. At each part we can get the new coordinates of unknown and assistant nodes. After exhaustive calculating all the data differences using equation (21), the ultimate position can always be acquired.

b: MAXIMUM GRADIENT DESCENT METHOD

The steps of the unconstrained maximum gradient descent method is shown as follows:

Step 1: the objective function $f(x^k)$

$$f(x^k) = \operatorname{argmin} \left(\sqrt{\sum_{i=1}^{m} \left((\bar{x}_i - x_i)^2 + (\bar{y}_i - y_i)^2 \right)} \right) \tag{22}$$

where x^0 respectively means the initial localizations of the unknown node and assistant nodes, meanwhile $x^k = (x_k, y_k)$ and localization estimation error $\varepsilon > 0\, k := 0$;

Step 2: calculating $\nabla f(x^k)$ based on equation (23)

$$\nabla f(x^k) = \begin{bmatrix} \dfrac{\partial f(x^k)}{\partial(x)} \\ \dfrac{\partial f(x^k)}{\partial(y)} \end{bmatrix} \tag{23}$$

If $\|\nabla f(x^k)\| \leq \varepsilon$, stopping the iteration and acquiring the output x^k, otherwise the process switches to Step 3.

Step 3: in order to acquire searching orientation p^k,

$$p^k = -\nabla f(x^k) = -\begin{bmatrix} \dfrac{\partial f(x^k)}{\partial(x)} \\ \dfrac{\partial f(x^k)}{\partial(y)} \end{bmatrix}$$

where $p^0 = -\nabla f(x^0)$, $k := 0$, $k := k + 1$.

Step 4: calculating step size t_k by formula (24)

$$f(x^k + t_k p^k) = \min_{t\geq 0} f(x^k + tp^k) \tag{24}$$

until $\min_{t\geq 0} f\left(x^k + tp^k\right)$ exiting, where $x^{k+1} = x^k + t_k p^k$, $k := k + 1$ and the process is switches to Step 2.

Step 5: optimal step size t_k resulting from equation (25)

$$f(x^k - t\nabla f(x^k)) = \varphi(t) \tag{25}$$

where $\dot{\varphi}(t) = 0$.

After finishing the whole schedule, the localization of the unknown node can be estimated with the help of the assistant nodes. Benefiting from the NLOS error mitigation algorithm, our system greatly improves the localization accuracy of the target, especially in complex indoor environments with severe NLOS error.

IV. EXPERIMENTAL EVALUATION

In order to evaluate our proposed system, several related experiments were conducted. The experiments' environment is shown as Fig. 10, and they have been conducted in Lab 306 of Yifu Technology Building of our university. As seen from Fig. 10, 12 RPs and 4 APs are set in the room, in addition many students are engaged in the daily work and many other experimental equipments are used in the lab including many Wi-Fi devices. Instead of general laptops, we use Beaglebone Black platforms with an Ubuntu system and omnidirectional antenna wireless cards to collect RSS data taking advantage of fast acquiring speeds and lower packet loss rate, which is shown in Fig. 8. Other equipment used in the experiments includes: a ProBook441s HP computer, 4 TP-LINK wireless routers named TL-WR742N.

FIGURE 8. RSS data collection platform.

To generate the TOF samples, we use the acknowledge mechanism of the IEEE 802.11 standard similar with [34]. Figure 9 is an illustration of this implementation mechanism.

For every data packet a round trip time is measured from the start of the transmission to the reception of the corresponding acknowledgement ($t_{MEAS}(d)$). The time the target waits is defined in the standard as t_{SIFS} and the duration of an acknowledgement t_{ACK} is constant. Therefore the propagation delay can be expressed as

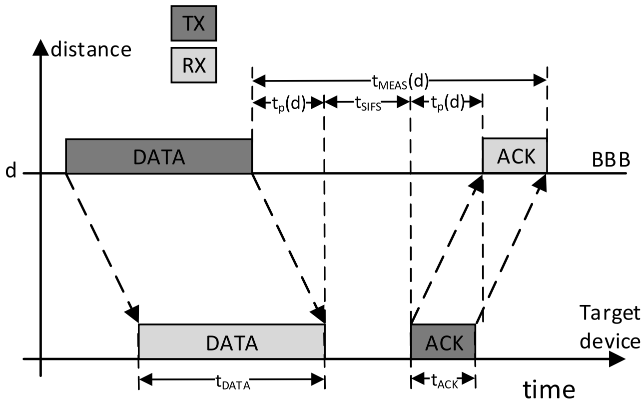


FIGURE 9. Illustration of TOF measurements.

$$t_P(d) = \frac{t_{MEAS}(d) - t_{SIFS} - t_{ACK}}{2}$$

From this equation, we can see that the propagation delay is proportional to the measured time for a given distance. In order to avoid additional delay, the time is measured directly in the Beaglebone Black firmware and not in the driver. This measurement is provided by the resolution of the General Purpose Timer of the firmware, which is clocked at 88MHz. Every time a measurement is made, the firmware writes the data into a defined register in the shared memory block which the driver can access to and retrieve the measurement once an acknowledgement is received.

The experiments were performed strictly following our system schedule. In the off-line fingerprint database construction stage, we select 10000 RSS data from the collected RSS data at each RP to reduce environmental influence and to improve the stability of RSS data as much as possible. During the localization stage, we collected RSS data of the target and surrounding nodes, performed initial localization, selected assistant nodes based on RSS time series similarity, measured distances between each pair of nodes, mitigated the NLOS ranging error and finally searched for the accurate position of the unknown node.

To evaluate the performance of our system in a complex indoor environment, we conducted the experiments three times, once during a time when nobody was in the lab, once during a time when 8 students were in the lab, and once when there were 30 students in the lab (in the latter two scenarios, the students worked as usual and moved from time to time). The NLOS environment grew worse as more and more students moved in the lab.

During the assistant nodes selection stage, the similarity threshold δ will affect the number of assistant nodes such that the larger it is, the fewer nodes will be selected. In the experiments, we studied how this threshold impacts the positioning accuracy in the 30 students' scenario, and the distribution is illustrated in Fig. 10.

We plot how the average localization error changes with the normalized threshold in Fig. 11. And it can be observed

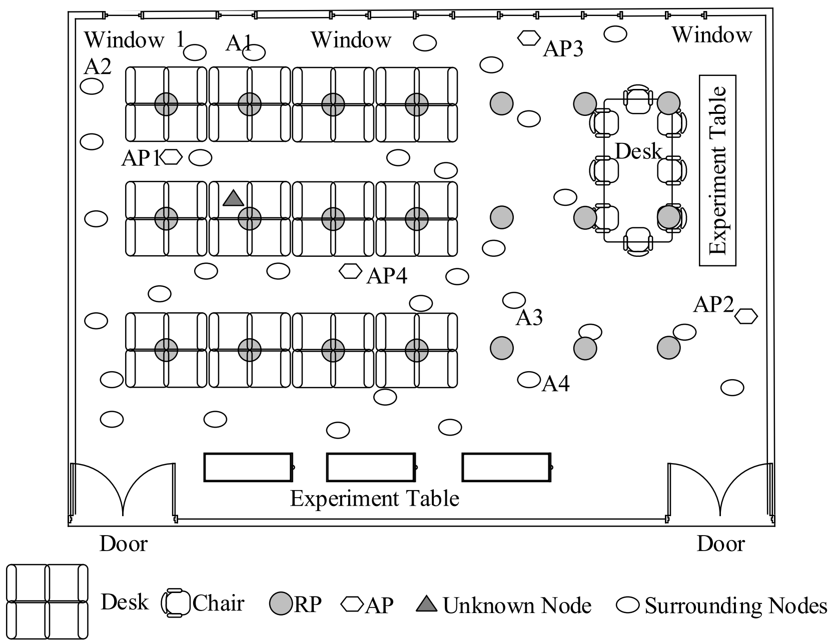


FIGURE 10. The students' distribution.

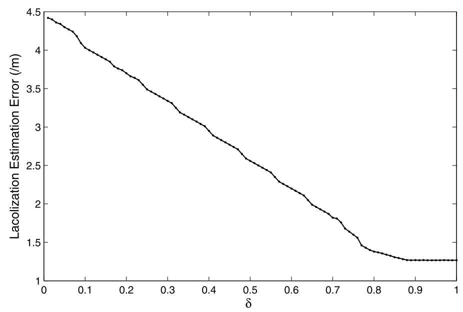


FIGURE 11. Impact of similarity threshold δ.

that when δ is 0.9, the localization error is almost stable with little fluctuation.

Furthermore, we also study the parameter of α described in the exhaustive algorithm of final searching stage mentioned in section 3.2.4. We also plot how the average localization error changes with α in Fig. 12. It can be observed that when α reaches 100, the localization estimation error almost remains the same.

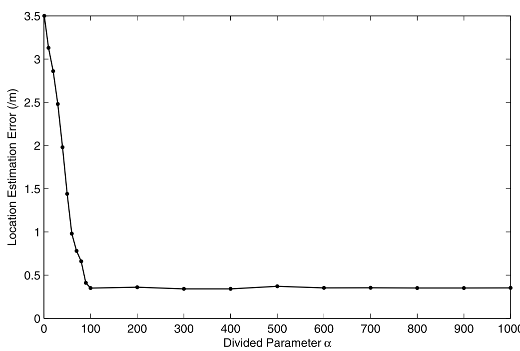


FIGURE 12. The divided parameter α.

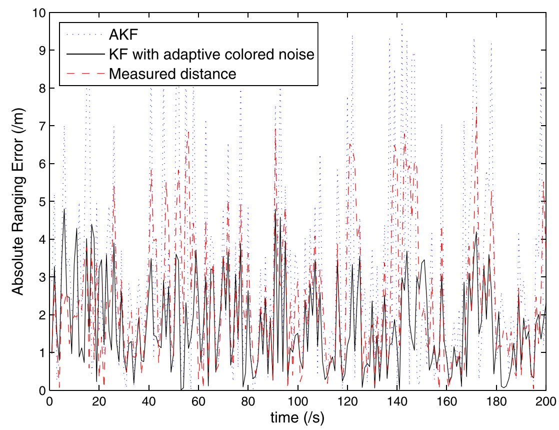


FIGURE 13. Absolute ranging error.

To evaluate the performance of the NLOS mitigation algorithm proposed in Section III, we simply measured the distance between node A1 and A4 (the real distance is 25.4 m) using a TOF-ranging method and later performed the Kalman Filter with adaptive colored noise elimination. As a comparison, we also implemented the adaptive Kalman filter (AKF) [35]. The NLOS mitigation performance is illustrated in Fig. 13 and 14.

As shown in Fig. 13, the error curve of the Kalman filter with adaptive colored noise is clearly located in the bottom, which indicates that the Kalman filter with adaptive colored noise in this paper has smaller error than the adaptive Kalman filter (AKF). Meanwhile, Fig. 14 demonstrates that the curve of the Kalman filter with adaptive colored noise lies in the top, and its cumulative probability is larger in the same absolute error. Moreover, the larger error of the Kalman filter with adaptive colored noise is smaller than the larger error of the

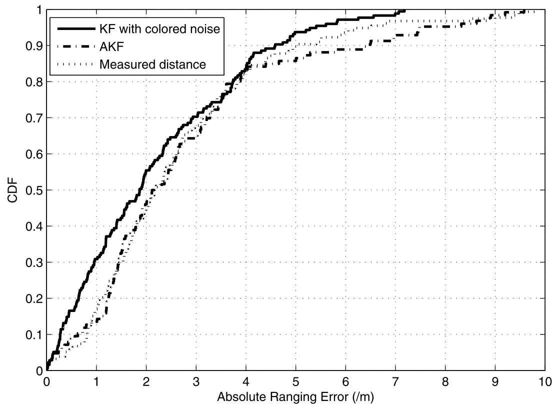


FIGURE 14. CDF of absolute ranging error.

measurements AKF. These results show that the Kalman filter with adaptive colored noise has better performance.

Based on the above parameters setting, we studied the overall performance of our Wi-Fi localization method based on Collaboration of Fingerprint and Assistant Nodes (dubbed CFAN) and compared it with other classical algorithms, such as RADAR [36] and Curve Filter with RSS [37] in the three scenarios (0, 8, and 30 students) mentioned above. Fig. 15, 16 and 17 show the localization error cumulative distribution functions (CDFs) of the testing points. Moreover, the performance line indicated by CFAN with NLOS presents the accuracy of localization system affected by NLOS error in a complex indoor environment.

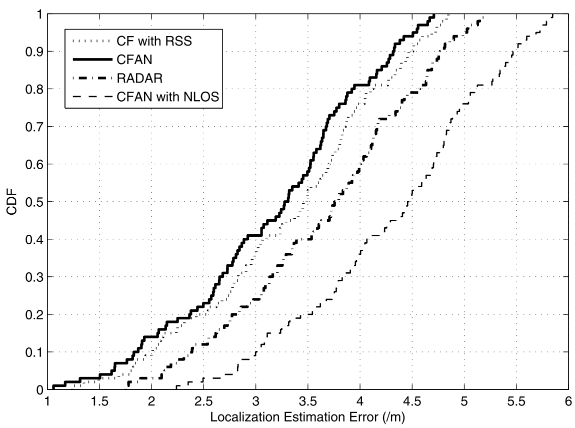


FIGURE 15. Localizaiton performance with 0 student.

From these figures, great performance improvements can be observed, with CDF curves shifting to the left significantly around nodes assistant and NLOS mitigation. It can be determined that NLOS mitigation is essential for improving localization accuracy: the long CDF tail has been reduced 30%.

Secondly, Fig. 15, 16 and 17 illustrate that the algorithm proposed in this paper has a lower localization estimation error and higher localization accuracy than other methods. Furthermore, the CFAN algorithm proposed in this paper reduces the maximum localization error from 9.1 m to 6.1 m over the experiments, as it eliminates NLOS.

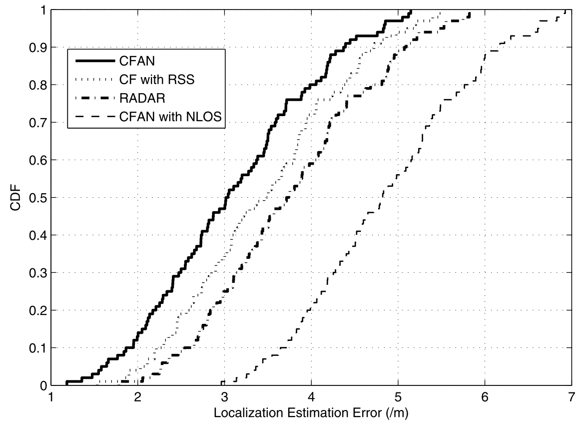


FIGURE 16. Localizaiton performance with 8 students.

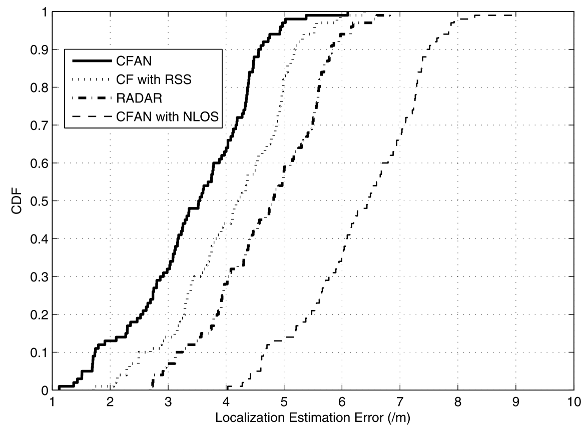


FIGURE 17. Localizaiton performance with 30 students.

From these figures, we can observe great performance improvements with our algorithm. As seen from Fig. 15, Fig. 16 and Fig. 17, NLOS severely affects the three algorithms in this paper with higher error, even though the algorithm proposed in this paper is better.

V. CONCLUSIONS

In complex indoor environments, multipath and NLOS conditions lead to large errors in measured values, thereby reducing indoor positioning accuracy. However, existing approaches have yet to be demonstrated to be effective in many business scenarios. To address the limitations of traditional indoor positioning methods, this paper presents a Wi-Fi indoor localization algorithm based on RSS and assistant nodes collaboration, which leverages positioning accuracy by TOF ranging and NLOS mitigation. Extensive experiments have demonstrated our approach is suitable for complex indoor environments. Our future work is to lessen the RSS data so that extra training effort can be effectively reduced. Moreover we will add more obstacles as in different real-time environment instead conducting experiment in limited environment.

REFERENCES

[1] K. Etemad and M.-Y. Lai, *Localization Based Service in WiMAX Networks*. New York, NY, USA: Wiley, 2010, pp. 345–360.

[2] S. He and S.-H. G. Chan, "Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 466–490, Jan. 2016.

[3] V. Moghtadaiee and A. G. Dempstera, "Indoor location fingerprinting using FM radio signals," *IEEE Trans. Broadcast.*, vol. 60, no. 2, pp. 336–346, Jun. 2014.

[4] H. Shin, Y. Chon, Y. Kim, and H. Cha, "A participatory service platform for indoor location-based services," *IEEE Pervasive Comput.*, vol. 14, no. 1, pp. 62–69, Jan./Mar. 2015.

[5] A. LaMarca and E. de Lara, *Location Systems: An Introduction to the Technology Behind Location Awareness*. San Rafael, CA, USA: Morgan & Claypool Press, 2008.

[6] L.-H. Chen, E. H.-K. Wu, M.-H. Jin, and G.-H. Chen, "Intelligent fusion of Wi-Fi and inertial sensor-based positioning systems for indoor pedestrian navigation," *IEEE Sensors J.*, vol. 14, no. 11, pp. 4034–4042, Nov. 2014.

[7] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, V. Handziski, and S. Sen, "A realistic evaluation and comparison of indoor location technologies: Experiences and lessons learned," in *Proc. 14th Int. Conf. Inf. Process. Sensor Netw.*, 2015, pp. 178–189.

[8] A. K. M. M. Hossain and W.-S. Soh, "A survey of calibration-free indoor positioning systems," *Comput. Commun.*, vol. 66, pp. 1–13 Jul. 2015.

[9] Y. Zhuang, Y. Li, H. Lan, Z. Syed, and N. El-Sheimy, "Smartphone-based WiFi access point localisation and propagation parameter estimation using crowdsourcing," *Electron. Lett.*, vol. 51, no. 17, pp. 1380–1382, Aug. 2015.

[10] H. Liu, J. Yang, S. Sidhom, Y. Wang, Y. Chen, and F. Ye, "Accurate WiFi based localization for smartphones using peer assistance," *IEEE Trans. Mobile Comput.*, vol. 13, no. 10, pp. 2199–2214, Oct. 2014.

[11] Z. Xiao, H. Wen, A. Markham, N. Trigoni, P. Blunsom, and J. Frolik, "Non-line-of-sight identification and mitigation using received signal strength," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1689–1702, Mar. 2015.

[12] J.-S. Leu, M.-C. Yu, and H.-J. Tzeng, "Improving indoor positioning precision by using received signal strength fingerprint and footprint based on weighted ambient Wi-Fi signals," *Comput. Netw.*, vol. 9, no. 1, pp. 329–340, 2015.

[13] M. Bshara, U. Orguner, F. Gustafsson, and L. Van Biesen, "Fingerprinting localization in wireless networks based on received-signal-strength measurements: A case study on WiMAX networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 283–294, Jan. 2010.

[14] Y. Heng, W. Y. Ya, and L. Bin, *Positioning Technology*. Beijing, China: Publishing House of Electronics Industry, 2013, pp. 162–164.

[15] P. Barsocchi, S. Lenzi, S. Chessa, and G. Giunta, "A novel approach to indoor RSSI localization by automatic calibration of the wireless propagation model," in *Proc. IEEE 69th VTC Spring*, Barcelona, Spain, Apr. 2009, pp. 1–5.

[16] B. Wang, G. Wu, S. Wang, and L. T. Yang, "Localization based on adaptive regulated neighborhood distance for wireless sensor networks with a general radio propagation model," *IEEE Sensors J.*, vol. 14, no. 11, pp. 3754–3762, Nov. 2014.

[17] R. Faragher and R. Harle, "Location fingerprinting with Bluetooth low energy beacons," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2418–2428, Nov. 2015.

[18] K. Kaemarungsi and P. Krishnamurthy, "Analysis of WLANŠs received signal strength indication for indoor location fingerprinting," *Pervasive Mobile Comput.*, vol. 8, no. 2, pp. 292–316, 2012.

[19] C.-H. Huang, L.-H. Lee, C. C. Ho, L.-L. Wu, and Z.-H. Lai, "Real-time RFID indoor positioning system based on Kalman-filter drift removal and Heron-bilateration location estimation," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 3, pp. 728–739, Mar. 2015.

[20] J. Niu, B. Wang, L. Shu, T. Q. Duong, and Y. Chen, "ZIL: An energy-efficient indoor localization system using ZigBee radio to detect WiFi fingerprints," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1431–1442, Jul. 2015.

[21] Y. Gao, J. Niu, R. Zhou, and G. Xing, "ZiFind: Exploiting cross-technology interference signatures for energy-efficient indoor localization," in *Proc. IEEE INFOCOM*, Turin, Italy, Apr. 2013, pp. 2940–2948.

[22] M. Shchekotov, "Indoor localization methods based on Wi-Fi lateration and signal strength data collection," in *Proc. 17th Conf. Open Innov. Assoc.*, Yaroslavl, Russia, 2015, pp. 186–191.

[23] S. Pagano, S. Peirani, and M. Valle, "Indoor ranging and localisation algorithm based on received signal strength indicator using statistic parameters for wireless sensor networks," *IET Wireless Sensor Syst.*, vol. 5, no. 5, pp. 243–249, 2015.

[24] C. Yang and H.-R. Shao, "WiFi-based indoor positioning," *IEEE Commun. Mag.*, vol. 3, no. 53, pp. 150–157, Mar. 2015.

[25] F. Zhao, H. Luo, H. Geng, and Q. Sun, "An RSSI gradient-based AP localization algorithm," *China Commun.*, vol. 11, no. 2, pp. 100–108, Feb. 2014.

[26] J. Jung, K. Kim, S. Yoo, M. Bae, and H. Kim, "RSSI localization with DB-assisted least error algorithm," in *Proc. 7th Int. Conf. Ubiquitous Future Netw.*, Sapporo, Japan, Jul. 2015, pp. 338–343.

[27] Y. Zhuang, Y. Li, H. Lan, Z. Syed, and N. El-Sheimy, "Wireless access point localization using nonlinear least squares and multi-level quality control," *IEEE Wireless Commun. Lett.*, vol. 4, no. 6, pp. 693–696, Dec. 2015.

[28] A. K. M. M. Hossain and W. Soh, "Cramer–Rao bound analysis of localization using signal strength difference as location fingerprint," in *Proc. IEEE INFOCOM*, San Diego, CA, USA, Mar. 2010, pp. 1–9.

[29] L. Banin, U. Schtzberg, and Y. Amizur, "Next generation indoor positioning system based on WiFi time of flight," in *Proc. Int. Tech. Meeting Satellite Division Inst. Navigat.*, Tampa, FL, USA, 2014, pp. 975–982.

[30] L. Qiyue, W. Zhong, L. Jie, S. Wei, and W. Jianping, "A novel adaptive Kalman filter based NLOS error mitigation algorithm," in *Proc. 17th IFAC Symp. Syst. Identificat. (SYSID)*, Beijing, China, 2015, pp. 1118–1123.

[31] L. Yang, H. Chen, Q. Cui, X. Fu, and Y. Zhang, "Probabilistic-KNN: A novel algorithm for passive indoor-localization scenario," in *Proc. IEEE 81st VTC Spring*, Glasgow, Scotland, May 2015, pp. 1–5.

[32] B. Jablonski, "Quaternion dynamic time warping," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1174–1183, Mar. 2012.

[33] C. K. Chui and G. Chen, *Kalman Filtering: With Real-Time Applications*. Beijing, China: Springer-Verlag, 2013.

[34] A. T. Mariakakis, S. Sen, J. Lee, and K.-H. Kim, "SAIL: Single access point-based indoor localization," in *Proc. 12th Annu. Int. Conf. Mobile Syst., Appl., Services*, 2014, pp. 315–328.

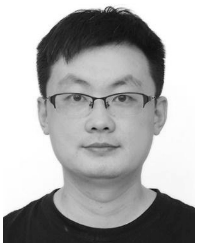
[35] J. He, Y. S. Geng, F. Liu, and C. Xu, "CC-KF: Enhanced TOA performance in multipath and NLOS indoor extreme environment," *IEEE Sensors J.*, vol. 14, no. 11, pp. 3766–3774, Nov. 2014.

[36] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM*, vol. 2. 2000, pp. 775–784.

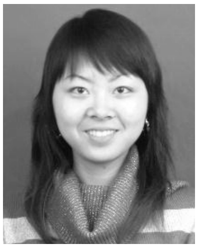
[37] B. Wang, S. Zhou, W. Liu, and Y. Mo, "Indoor localization based on curve fitting and location search using received signal strength," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 572–582, Jan. 2015.

QIYUE LI (M'13) received the B.E. degree in electronics engineering from Wuhan University, China, in 2003, and the Ph.D. degree in communication and information system from the University of Science and Technology, China. He is currently an Associate Professor with the Hefei University of Technology, Anhui, China. From 2008 to 2011, he did post-doctoral research with the School of Computer Science and Technology, University of Science and Technology. His research interest includes wireless sensor networks and indoor localization using wireless networks. He is a member of the IEICE.

WEI LI received the bachelor's degree in automation from the Hefei University of Technology, China, in 2014, where he is currently pursuing the master's degree in control engineering. His research interest includes wireless sensor networks and indoor localization using wireless networks.



WEI SUN received the B.E. degree in automation, the M.E. degree in detection technology and automatic equipment, and the Ph.D. degree in electrical engineering from the Hefei University of Technology, China, in 2004, 2007, and 2012, respectively, where he is currently a Lecturer. His research interest includes wireless sensor networks and smart grid.



JIE LI received the B.E. degree in electronics engineering from University of Science and Technology, China, in 2006, and the Ph.D. degree from the University of Science and Technology, China, in 2011. She is currently an Associate Professor with the Hefei University of Technology, Anhui, China. From 2012 to 2012, she was a Research Assistant with the National Institute of Informatics, Japan. Her research interest includes cognitive radio network, scalable video multicast, and indoor localization using wireless networks.

ZHI LIU (S'11–M'14) received the B.E. degree in computer science and technology from the University of Science and Technology of China, China, in 2009, and the Ph.D. degree in informatics from the National Institute of Informatics, The Graduate University for Advanced Studies (Sokendai), Tokyo, Japan. He is currently a Junior Researcher (Assistant Professor) with Waseda University, Tokyo. He was a JSPS Research Fellow with the National Institute of Informatics, The Graduate University for Advanced Studies (Sokendai) from 2012 to 2014. From 2009 to 2014, he was a Research Assistant with the National Institute of Informatics, The Graduate University for Advanced Studies (Sokendai).

His research interest includes wireless networks, video/image processing, and transmission. He was a recipient of the IEEE StreamComm2011 Best Student Paper Award, the VTC2014-Spring Young Researchers Encouragement Award, and the 2015 IEICE Young Researchers Award. He has been a Guest Editor of *Sensors* and the *IEICE Transactions on Information and Systems*. He has been serving as the TPC Chair of the 2017 IEEE Workshop on Game Theory in Computer Communications (in conjunction with the IEEE Consumer Communications and Networking Conference). He is a member of the IEICE.