Received: 18 August 2017 | Revised: 18 January 2019 | Accepted: 13 February 2019

DOI: 10.4218/etrij.2017-0116

ORIGINAL ARTICLE

WILEY ETRI Journal

Multiregional secure localization using compressive sensing in wireless sensor networks

Chang Liu | **Xiangju Yao** | **Juan Luo**

College of Computer Science and Electronic Engineering, Hunan University, Changsha, China

Correspondence
Juan Luo, College of Computer Science and Electronic Engineering, Hunan University, Changsha, China.
Email: juanluo@hnu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 61672220; Key Technology Research and Development Plan of Hunan, Grant/Award Number: 2017GK2030.

Security and accuracy are two issues in the localization of wireless sensor networks (WSNs) that are difficult to balance in hostile indoor environments. Massive numbers of malicious positioning requests may cause the functional failure of an entire WSN. To eliminate the misjudgments caused by malicious nodes, we propose a compressive-sensing–based multiregional secure localization (CSMR_SL) algorithm to reduce the impact of malicious users on secure positioning by considering the resource-constrained nature of WSNs. In CSMR_SL, a multiregion offline mechanism is introduced to identify malicious nodes and a preprocessing procedure is adopted to weight and balance the contributions of anchor nodes. Simulation results show that CSMR_SL may significantly improve robustness against attacks and reduce the influence of indoor environments while maintaining sufficient accuracy levels.

KEYWORDS
compressive sensing, indoor localization, received signal strength, secure localization, wireless networks

1 | INTRODUCTION

Advanced technologies, including advanced sensing, embedded computing, and short range communications, have been introduced into sensors to extend the applications of indoor localization systems by providing fast and precise positioning for wireless sensor networks (WSNs). In recent years, indoor localization systems [1] have been successfully implemented in a variety of applications, including major museum navigation, warehouse inventory inquiry, and parking garage management [2,3].

As sensors become smaller and more power efficient, it is unrealistic to equip all sensors with global positioning systems. The accuracy requirements of indoor localization systems are strict and have become a key factor because nearly all indoor applications depend on localization information. According to [4], indoor localization methods can be categorized into three main types: time of arrival (TOA), time difference of arrival (TDOA), and received signal strength (RSS) methods. TOA localization requires synchronization between a source and receivers. The authors of [5] proposed a base station selection scheme to reduce localization error in real indoor experiments based on TOA localization. Although TDOA does not require synchronization, it does require additional hardware devices [6]. RSS localization can achieve significant benefits with the easiest implementation and lowest cost among the three positioning technologies. However, RSS is susceptible to indoor environmental changes. Wireless signal propagation during indoor communications is difficult to capture and characterize based on environmental effects on signal absorption, attenuation, reflection, or a combination of these factors [7]. Therefore, when indoor environments are intricate, the performance of RSS-based algorithms tends to deteriorate. The above algorithms all suffer from low localization accuracy, high energy consumption, unnecessary wireless communication for redundant signal collection, and per-sensor computation of signal readings.

Significant research efforts have been devoted to exploring different in-network localization solutions for WSNs based on wireless signal readings. Fingerprint localization

catches have recently received significant attention based on their easy implementation [8–10]. Bahl and Padmanabhan proposed a classic indoor localization mechanism called fingerprint localization in [9]. Signals are collected and processed using the Nyquist sampling principle (double signal maximum frequency) in fingerprint approaches based on empirical measurements with signal propagation modeling. Locations are then identified based on the overlapping coverage in the region of interest. Based on the use of an offline signal dictionary, the size of signal samples in fingerprint localization is sufficient for reflecting complex indoor environmental factors, meaning this method can reduce localization errors efficiently. However, most existing fingerprint localization methods assume that an indoor environment is constant and rarely changes, which is not true in most real-world scenarios. In changing environments, localization using the fingerprint method may produce large communication overhead and high computational costs during signal collection and processing [8,10].

To reduce large communication overhead and high computational costs, the compressive sensing (CS) theory [11] was proposed to recover sparse signals using far fewer noisy measurements compared to the number predicted by the Nyquist theorem. Based on the development of this theory in recent years, Liu et al. [12] designed a multiple-source localization model based on compressive sensing for WSNs. This model can optimize redundant dictionary arrangements with excellent performance for multiple source localization. In [13], a sampling database was generated in an offline phase by collecting signals from individual network grids. A clustering algorithm was then used to group all anchors and construct a redundant dictionary. In the online phase, a cluster header was chosen to determine the most likely cluster for an unknown node and CS was performed to recover the original signal and derive the position of the unknown node.

The above algorithms largely focus on accurate localization without consideration for hostile indoor environments. Consequently, secure localization schemes have attracted significant attention in recent years [4,14]. Secure distance-based localization was investigated in [14] and bounded localization error was formalized for secure and robust distance-based localization. Another secure localization algorithm was proposed in [4] with the goal of withstanding attacks from malicious users. By using an iterative gradient descent approach, the localization accuracy of the algorithm was improved significantly, even in hostile indoor environments. A secure fingerprint localization (SFL) method based on a trusted RSS database was proposed in [15]. This method is robust to variable environments, impaired access points, and the introduction of new access points. The SFL method updates a fingerprint database according to a reference anchor node voting strategy that can reduce the interference caused by user measurement data and abnormal RSS messages.

Hostile indoor environments are a major constraint for reliable indoor localization. The confidentiality of localization is essential in certain environments, such as shopping malls, car parks, hotels, and cafes. It is very easy for malicious users to capture location information and send false information in overcrowded public locations. Most indoor localization methods in the literature [8,12,16,17] assume that no anchor nodes, which have predetermined locations, act as malicious nodes to prevent the accurate localization of target nodes. Because target nodes are compromised by malicious sensor nodes, position estimates can be adversely affected without effective and secure localization mechanisms to eliminate the interference of incorrect location measurements. Therefore, it is reasonable to consider the presence of malicious nodes as a primary performance metric.

In this paper, we propose a multiregion-based secure localization method using the compressive sensing theory (CSMR_SL). The proposed method consists of both offline and online operations. We aim to dynamically detect and eliminate the effects of incorrect location references and construct an adaptive redundant dictionary with reliable signal measurements for CS operation. Therefore, CSMR_SL can achieve a high level of localization accuracy with low energy consumption and few RSS samples in hostile indoor environments. The main contributions of this work can be summarized as follows:

1. We design a multiregion mechanism with a transmission-range-based secret key to detect and eliminate malicious anchors dynamically in hostile indoor environments.
2. We present a preprocessing procedure for the online phase to weight RSS measurements and balance the contribution of each anchor node.
3. We introduce CS theory in a secure localization algorithm to recover a target location from dynamic RSS measurements in the online phase to reduce misjudgments caused by malicious nodes.
4. We propose the CSMR_SL algorithm, which is assisted by RSS, to increase localization accuracy and minimize energy consumption.

The remainder of this paper is organized as follows. Section 2 introduces the fundamentals of secure localization system architectures based on CS. The proposed CSMR_SL algorithm is described in Section 3 and a performance evaluation of the proposed algorithm is presented in Section 4. Finally, Section 5 concludes this paper and discusses our plans for future work.

2 | CS-BASED SECURE LOCALIZATION SYSTEM

A method combining an efficient CS localization algorithm with a dynamic security key verification mechanism could effectively solve the secure localization problem. In this section, we describe indoor localization scenarios and present an overview of the proposed secure localization architecture based on CS. For convenience, the notations used in this paper are summarized in Table 1.

2.1 | Indoor localization scenario

Figure 1 presents an indoor localization scenario. Anchor nodes know their own positions and can communicate with neighboring nodes through broadcasting. The position of a target node can be estimated based on the relevant RSS information from its perceived neighboring anchor nodes. The anchor nodes are deployed in an equilateral triangle layout in this scenario. When a location area is fixed, the entire area can be covered by a minimal number of anchor nodes. In this manner, hardware costs can be reduced. The authors of [18] adopted an identical scenario for testing their algorithms. However, anchor nodes can be seized and transformed into malicious anchor nodes in certain unfriendly environments. A malicious anchor node is considered as a type of information interference node in this setting. Consider a scenario in which a target node (unknown position) enters a wireless network deployed indoors. To locate the target node, adjacent anchors (known positions), which can be perceived by the target node, send messages to the target node. These messages contain the coordinates of the anchors and RSS readings. The target then sends the received information to a base station (or server). Based on the interference of objects (such as the movement of people) and impact of multipath fading, the target may send large amounts of new information to the base station periodically. CS theory, also referred to as compressed sampling/compressed sensing, is based on the sparse characteristics of such signals, which can be used to obtain discrete signal samples under sampling rates far lower than the Nyquist sampling rate. A perfect reconstruction signal can then be obtained by using a nonlinear reconstruction algorithm. Most studies on CS have focused on image processing. The original CS localization algorithm generates a sample library during the offline phase. Through node selection, it then selects a small number of highly correlated samples and uses a signal recovery algorithm for positioning. Compared to the traditional RSS-based fingerprint database location algorithm, it significantly reduces the cost of positioning. Overall, CS is designed to enhance localization efficiency.

TABLE 1 Notations used in this paper

Symbol	Description
N	Number of samples
$\boldsymbol{\Psi}$	A measurement matrix
$\mathbf{K}$	A constant number representing sparsity
M	Number of measurements
$\mathbf{P}$	M-dimensional projection under a measurement matrix
$\boldsymbol{\Phi}$	Sampling matrix
$\tilde{N}$	A subset of N
Θ	Recovery signal
ε	Error of recovery
$\mathbf{R}$	Relationship matrix
r	Anchor communication radius
r_1	The length of equilateral triangle sides
L	Number of anchors
δ	Location error

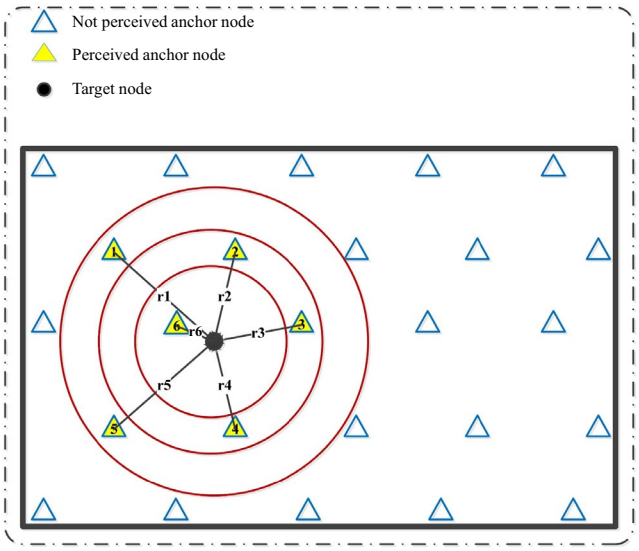

FIGURE 1 Example indoor localization scenario

2.2 | System architecture

The system architecture of CSMR_SL is presented in Figure 2. It consists of an offline phase, secure monitoring phase, and online phase. The core task of the offline phase is to construct a complete database. In the secure monitoring phase, which uses a key-checking mechanism, the location information of malicious anchor nodes is filtered. The online phase then takes the trusted complete offline database as an input to construct a dynamic matrix with reliable measurements for secure positioning.

(a) *Offline Phase*: Establishing an RSS fingerprint database is the main task of the offline phase. The RSS fingerprint database records RSS information transmitted from all

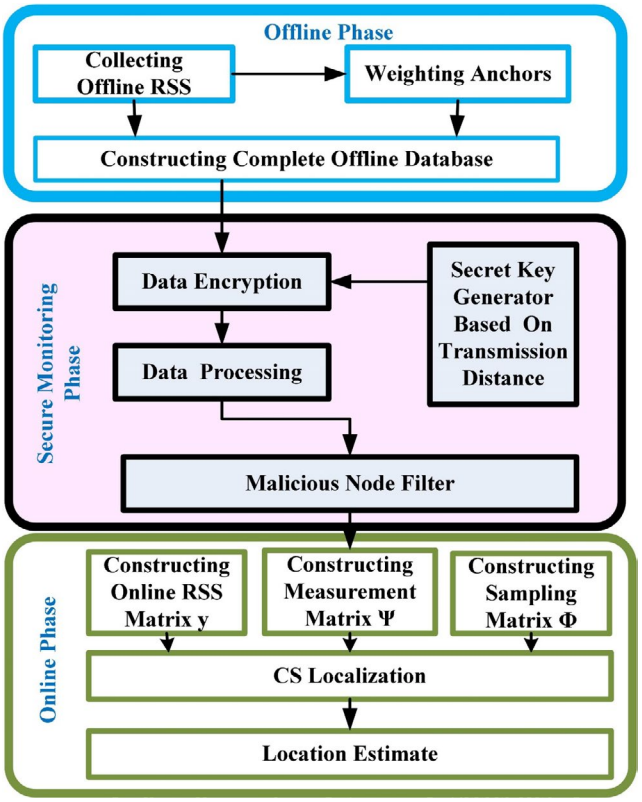

FIGURE 2 Architecture of CSMR_SL

anchor nodes in each reference grid, as well as the weights of each anchor node, corresponding locations, and total weight value of all anchor nodes.

The indoor locating area is divided into several reference grids of equal size to collect offline RSS information. Anchor nodes are laid out in the form of equilateral triangles within the locating area and neighboring nodes can communicate with each other. The reference node in each grid receives information from neighboring anchors in its communication range and transmits the information to the base station, where the offline fingerprint database is recorded.

A weighting mechanism is used to obtain each grid's weight based on offline RSS information. Additionally, these weight values can be used to choose reliable anchor nodes for locating targets.

(b) *Secure Monitoring Phase*: The secret key generator and malicious node filter are important parts of the secure monitoring phase. In contrast to traditional key distribution authentication, the preinstallation of secret keys is not shared across the entire network.

Each perceived anchor node has a unique key based on its anchor ID and the transmission distance between the target node and perceived anchor node. The target node may encrypt a secret key and transmit it only to the corresponding anchor node, meaning other perceived anchors will be unable to decipher the secret key. By using this key-checking mechanism, malicious anchor nodes can be excluded. Therefore, the estimated location of the target node is reliable and accurate after the secure monitoring phase.

(c) *Online Phase*: A target node enters the locating area and must be positioned. A relationship matrix R is constructed by adopting an overlapping mechanism. This matrix is the key to constructing a sampling matrix and online RSS matrix for CS. After obtaining R, CS generates a sampling matrix $\mathbf{\Phi}$ and measurement matrix $\mathbf{\Psi}$, and collects online RSS readings y. Equation (1) is used by CS to recover a signal θ, which includes location information.

$$y = \mathbf{\Phi}\Theta + \varepsilon = \mathbf{\Phi}\mathbf{\Psi}\theta + \varepsilon, \tag{1}$$

where Θ is a column vector denoting a discrete-time signal in the N-dimensional space and ε is the error of recovery.

3 | CS-BASED MULTIREGIONAL SECURE LOCALIZATION

To introduce the model of CS localization into WSNs, we propose a mutual authentication mechanism that is used to construct a trusted measurement matrix $\mathbf{\Psi}$. Based on this matrix, a noniterative transmission-distance-based mutual authentication scheme is used to detect the interference of incorrect location measurements.

We assume that a localization area is divided into $N \times N$ grids and that L anchors are deployed in the shape of an equilateral triangle, as shown in Figure 1. We randomly place a target node in the area for location. In this setting, "randomly" indicates that the location of the target node is different for each trial. The radius of an anchor's communication range is r and RSS technology is adopted to weight anchors.

The CSMR_SL algorithm consists of the following stages:

3.1 | Weight-based offline approach

We must collect RSS information from the entire positioning area to construct a complete RSS fingerprint database. Suppose that the total number of RSS samples that a grid has received from anchor i is q. Then, the average RSS value for this specific anchor in this grid is defined as follows:

$$\varphi_{i,j} = \sum_{\tau=1}^{q} \varphi_{i,j}(\tau)/q, \tag{2}$$

where, $\varphi_{i,j}(\tau)$ represents the τth RSS that the jth grid received from anchor I and $\varphi_{i,j}$ represents the average RSS value from a specific anchor i for grid j.

The total number of anchor nodes is L. An original database matrix $\mathbf{\Psi}$ is defined as follows:

$$\mathbf{\Psi} = \begin{Bmatrix} \varphi_{1,1} & \varphi_{1,2} & \cdots & \varphi_{1,N} \\ \varphi_{2,1} & \varphi_{2,2} & \cdots & \varphi_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{L,1} & \varphi_{L,2} & \cdots & \varphi_{L,N} \end{Bmatrix}. \tag{3}$$

If a particular anchor (anchor and anchor node are used interchangeably) is not perceived by a reference anchor, then $\varphi_{i,j} = 0$. We formalize this rule as the weighted value of an anchor node, which reflects its contribution to the entire offline sampling phase. In the complete offline database, w_i denotes the jth anchor's weighted value.

In this work, we employ an anchor's utilization rate to estimate its weighted value. A reference anchor, which receives a large amount of RSS information from adjacent anchors, is assigned a large weight value. Therefore, the weight values are directly proportional to the sensing of adjacent anchors. We define a weight value as follows:

$$W = \left[w_1, w_2, ..., w_L\right]^T \tag{4}$$

and

$$w_i = \left(\sum_{k=1}^{N} \mathrm{Num}_k\right) \Big/ N, \tag{5}$$

where w_i describes the jth anchor's weighted value, Num_k is the number of grids in which the anchor is perceived, and N is the total number of grids.

Because we conducted experiments in two dimensions, the most important parameters for localization are the x and y coordinates of the anchors. Each anchor's x coordinate, y coordinate, RSS value, and weight value are combined to form a complete offline database, as shown in (6).

$$(x_i, y_i; \psi_i, \Delta_i; w_i), \quad i = 1, 2, ..., N, \tag{6}$$

where (x_i, y_i) is the position of anchor i.

3.2 | Security detection mechanism

(a) *Establish a security key:* $D = \{r_1, r_2, r_3\}$ are the radius of the receiving ranges of three different anchors. The deployment of anchor nodes is presented in Figure 1. Therefore, the distance between any two adjacent anchor nodes is a fixed value denoted as r_1. r_3 is the communication radius of a target node that can receive RSS information within the maximum range of the anchors. r_2 is the average of r_1 and r_3. The measured RSS values of the three positions, denoted $T = \{t_1, t_2, t_3\}$, are received as three different thresholds. Based on the transmission radius value D and threshold value $\mathbf{\Psi}'$, the following random number matrix for an anchor node is received by the target node:

$$\mathrm{Num} = \begin{bmatrix} \mathrm{num}_{1,1} & \mathrm{num}_{1,2} & \cdots & \mathrm{num}_{1,n} \\ \mathrm{num}_{2,1} & \mathrm{num}_{2,2} & \cdots & \mathrm{num}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{num}_{L,1} & \mathrm{num}_{L,2} & \cdots & \mathrm{num}_{L,n} \end{bmatrix}, \tag{7}$$

where $\mathbf{\Psi}'$ is the online RSS matrix and $\varphi'_{i,j}$ is an element of $\mathbf{\Psi}'$. $\mathrm{num}_{i,j}$ indicates the random number sent by the jth anchor node to the ith target node. When $\varphi'_{i,j}$ is between the thresholds of t_k and t_{k+1}, the corresponding transmission radius of the jth anchor node is r_{k+1}, where k is one or two. When the distance between the target node and anchor node is greater than r_3, the RSS received by the target node is set to zero. An anchor can only sense nearby target nodes. For any other nodes, $\mathrm{num}_{i,j}$ is zero. n represents the total number of target nodes.

(b) *Detection and exclusion of malicious anchor nodes*: We use $\mathbf{\Psi}'$ to verify the random numbers in Num to determine if they are consistent with the random number generation algorithm of reversible computing. If $\varphi'_{i,j}$ and $\mathrm{num}_{i,j}$ are not satisfied, then the location information of the ith anchor node will not be used during the localization of the jth target node.

3.3 | Multiregion mutual authentication mechanism (MRMAM)

After constructing a complete offline database, a complex indoor environment can be described exactly. In this subsection, we first propose a multiregion algorithm to construct a relationship matrix to improve security.

We formalize relationship matrix construction as a clustering problem. The MRMAM is proposed to solve this problem. The objective of the MRMAM is to reduce the impact of malicious users on secure positioning and minimize the target region by adopting an overlapping scheme.

The basic idea of the overlapping scheme is to overlap the communication regions of all candidate anchors, which are chosen by computing the fitness functions of each anchor. In our overlapping scheme, fitter anchors have higher online RSS values and offline weight values. Therefore, we simply define the fitness function f as the weighted online RSS as follows:

$$f = \mathbf{\Psi}' \times \mathbf{W}, \tag{8}$$

where $\mathbf{\Psi}'$ is the online RSS matrix and $\mathbf{W}$ is the offline weight value matrix.

We select weight values from the offline database and multiply each anchor's online RSS by the corresponding weight value. If an anchor cannot hear from the target, then its function f will be equal to zero. The top nonzero anchor communication regions are then overlapped and the grids in the overlapping region are identified. These grids have a special relationship with the unknown node. We formulate the relationship matrix $\mathbf{R}$ as a vector. The number of nonzero values in $\mathbf{R}$ is denoted as $\tilde{N}$.

Because grids have special relationships with the unknown node, R can be defined as follows:

$$\mathbf{R} = \left[R_1, R_2, R_3, ..., R_\varsigma, ..., R_{\tilde{N}}\right], \quad R_\varsigma \in \{0, 1\}. \tag{9}$$

As shown in Figure 3, a target node T enters the locating region. The target node T can communicate with anchor nodes 1, 2, 3, and 4. We overlap these four anchor node

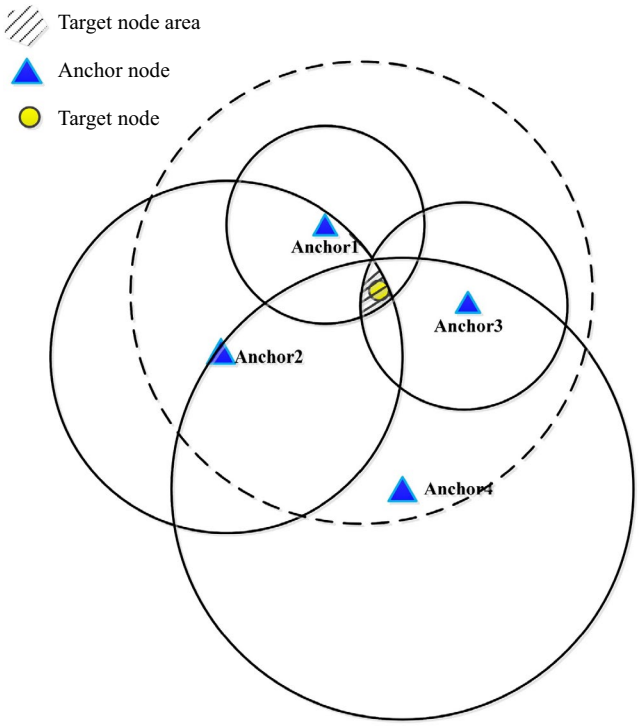

FIGURE 3 Overlapping scheme

communication ranges based on their fitness functions f to derive an overlapping area that contains the target node. The area represents a rough location estimate for the target. In this case, the number of nonzero values in the relationship matrix $\mathbf{R}$ is four (ie, $\tilde{N} = 4$).

The relationship matrix $\mathbf{R}$ is the foundation of the entire online localization phase, which identifies the locating region. We can construct a sampling matrix for localization based on $\mathbf{R}$.

3.4 | CS online approach

CS supports localization in which sparse signals are constructed periodically. This model fits many scenarios in which data are collected from sensors at a rate lower than the Nyquist rate [19].

CS localization depends on two main properties: sparsity and incoherence. Sparsity indicates that natural signals in Θ are sparse when expressed in the following convenient form:

$$\Theta = \mathbf{\Psi}\theta = \sum_{i=1}^{N} \mathbf{\Psi}(i)\theta_i. \tag{10}$$

A signal Θ is K-sparse if its transformation coefficient is K on the transformation basis $\mathbf{\Psi}$. θ represents a sparse signal. Instead of requiring N samples, only a small set of measurements is required to recover θ via CS.

Incoherence between $\mathbf{\Phi}$ and $\mathbf{\Psi}$ is another important property to ensure successful recovery of an original signal via CS. Candès et al. proved that a signal can be recovered only if $\mathbf{\Phi}$ and $\mathbf{\Psi}$ meet the restricted isometry property (RIP) condition [11,20–22]. The RIP is widely used to judge incoherence. In [23], it was proven that the RIP can be perfectly satisfied when

$$M \geq O(K \log(\tilde{N}/K)). \tag{11}$$

When measurements satisfy (11), a sparse signal θ can be recovered by using CS localization. Figure 4 presents the relationship between sparsity K, the number of chosen anchors $\tilde{N}$, and number of measurements M. $\tilde{N}$ is the dimension of the signal θ, which is equal to the number of anchors chosen for locating the target node. These results reveal that $\tilde{N}$ grows rapidly with an increase in sparsity for the same number of measurements. Therefore, we can conclude that K is inversely proportional to $\tilde{N}$. As mentioned above, a large value of K can make CS more effective, but a small value of $\tilde{N}$ can make signal recovery excessively complex. Therefore, in practice, one should choose an appropriate tradeoff between $\tilde{N}$ and K.

Additionally, according to (11), $\tilde{N}$ depends on the size of $\mathbf{\Phi}$ and $\mathbf{\Psi}$. Therefore, we should minimize the size of $\mathbf{\Phi}$ and $\mathbf{\Psi}$ for a given sparsity K. In [24], a clustering algorithm was utilized to minimize the size of $\tilde{N}$, but this process resulted in significant computational complexity. However, because CSMR_SL adopts the relationship matrix $\mathbf{R}$, it can effectively minimize the size of $\mathbf{\Phi}$ and $\mathbf{\Psi}$ with lower computational complexity. Following CS localization, we can recover the signal θ. Because the recovered signal cannot locate the target node exactly, it is necessary to estimate the final accurate location of the target by performing weight calculation again.

(a) *Constructing the sampling matrix*: The sampling matrix $\mathbf{\Psi}$ is the sparse basis on which the measured signals have a sparsity coefficient K. We construct $\mathbf{\Psi}$ as a selection problem. Useful information from the offline database should be

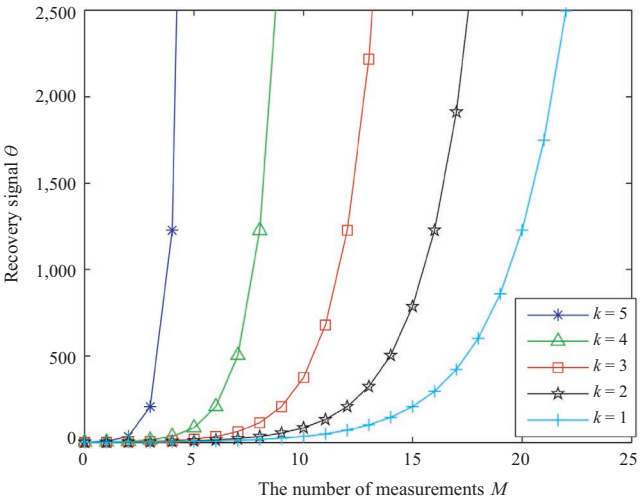

FIGURE 4 Relationship between M and $\tilde{N}$ for different values of K

selected according to **R**. The anchors whose data in **R** are nonzero are considered to be the most useful anchors. The RSS information of these anchors is selected from the offline database to form the sampling matrix **Ψ**. Because the total number of nonzero anchors in **R** is $\tilde{N}$, **Ψ** is a matrix, meaning the sparsity coefficient K depends on $\tilde{N}$.

(b) *Constructing the measurement matrix*: The measurement matrix **Φ** is formed from the online perceived anchors. Locating a target does not require all anchor measurements from the entire region. If we use all RSS readings from all grids, large computational cost is incurred and accuracy is reduced. Therefore, we select a smaller group of strong anchors for location. The number of chosen anchors is denoted as M. In the offline phase, all anchors have their own sequence number. Each row of **Φ** represents the sequence number of an anchor and a value of one indicates that an anchor's sequence number is the same as the row number. Therefore, the matrix **Φ** is defined as follows:

$$\mathbf{\Phi}=\begin{Bmatrix} \phi_{1,1} & \phi_{1,2} & \cdots & \phi_{1,M} \\ \phi_{2,1} & \phi_{2,2} & \cdots & \phi_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{L,1} & \phi_{L,2} & \cdots & \phi_{L,M} \end{Bmatrix}, \quad \phi_{i,j}\in 0,1. \tag{12}$$

If the sequence number of anchor i is j, then $\phi_{i,j}=1$.

(c) *CS localization*: P is an M-dimensional projective under the measurement matrix **Φ**. If **Φ** and **Ψ** satisfy (11), then θ can be recovered. We select the online RSS readings of M anchors from P. It is well known that a sparse signal θ can be recovered by using convex optimization, greedy optimization, or statistical optimization. In this work, the l_1-norm minimization algorithm is used to recover θ. This problem can be formulated as follows [25]:

$$\theta=\arg\min_{\theta\in R^{\tilde{N}}}|\theta|_1, \quad \text{such that } \mathbf{P}=\mathbf{\Psi\Phi}\theta+\varepsilon. \tag{13}$$

(d) *Localization estimate*: After recovering the sparse signal θ using l_1-norm minimization, an ideal 0–1 matrix can be constructed. In conventional wireless scheduling approaches, the highest value in θ is often taken as the target's final location grid. However, this is not necessarily an accurate location for the target. Nonzero values in θ must be considered to locate the target accurately. Therefore, we employ a weighting mechanism to optimize the target's final coordinates. θ is considered to represent the probabilities that different grids contain an unknown node. Therefore, θ and the corresponding grid coordinates are weighted to guarantee the accuracy of an unknown node's coordinates. This process is defined as follows:

$$P_x=\sum_{k=1}^{\tilde{N}}\theta_k x_k \tag{14}$$

and

$$P_y=\sum_{k=1}^{\tilde{N}}\theta_k x_k, \tag{15}$$

where P_x and P_y are target's x coordinate and y coordinate, respectively.

Algorithm 1 contains the pseudocode for the CSMR_SL algorithm. This algorithm requires $O(\tilde{N})$ time to locate a target. The base station in a wireless network is the key component of a secure positioning system. The storage of the offline database and the calculation and execution of the entire algorithm take place at the base station.

Algorithm 1 CSMR_SL Algorithm

Input: $RSS(r_i)$, **Ψ**′, **W**, r_1, r.

Output: Target position (P_x, P_y).

/*Initialization*/

1: $\theta=0$, meaning the set of nonzero elements in θ is null.

/*Steps*/

2: Formulate the region deployment $\{r_n\}$ as

$$r_k=r_1+(k-1)*d, \quad d=\frac{r_n-r_1}{n-1};$$

3: Set $d(i,j)$ to the distance between the target node j and anchor node i;

4: **if** $r_{k-1}\le d(i,j)<r_k$ **then**

5: $d(i,j)=r_k$;

6: **end if**

7: Set anchors to the corresponding region $S_i(r_i)$;

8: $\text{Num}(r_i)\leftarrow \text{Random}(0,1)+r_i$;

9: Generate a random symmetric key $K_m=\{k_1, k_2, ..., k_i, .., k_n\}$ for the target node m as follows:

$k_i=\{S_i(r_i), m, \text{Num}(r_i), \text{RSS}(r_i)\}, \quad i\in(1,n)$;

10: Set the anchor node $a\in S_i(r_i)$;

11: **if** The key value of anchor node a does not equal k_i **then**

12: Anchor node a is a malicious node;

13: Eliminate anchor node a from $S_i(r_i)$;

14: **else**

15: Monitor all anchor nodes;

16: **end if**

17: Construct the sampling matrix **Ψ** and observation matrix **Φ**;

18: Apply the CS localization algorithm to calculate a rough location as follows:

$$\theta=\arg\min_{\theta\in R^{\tilde{N}}}\|\theta\|_l, \text{ such that } y=\mathbf{\Psi\Phi}\theta+\varepsilon;$$

19: Calculate the target's position p_x, p_y as

$$P_x=\sum_{k=1}^{\tilde{N}}\theta_k x_k$$

and

$$P_y=\sum_{k=1}^{\tilde{N}}\theta_k x_k$$

until the end of the iteration on θ;

20: return the target position (P_x, P_y).

4 | SIMULATION RESULTS

We conducted simulation experiments using MATLAB with 17 anchors (known positions) and various targets (unknown positions) that were uniformly and independently distributed

in a 100 × 100-m^2 region. To fully evaluate the performance of CSMR_SL, we compare it to a traditional localization algorithm (CS_NSL) based on CS theory [12] with no protection against malicious nodes and a secure localization algorithm (CS_SL) whose secret key is shared across the entire network. To some extent, CS_SL ensures the security of locations, but neglects the accuracy of locations. In contrast, CS_NSL ensures the accuracy of positioning, but neglects the safety of the environment. In this paper, forgery attacks and replay attacks [26] are discussed to evaluate their influence on secure localization algorithms.

All nodes (including anchors and targets) operate on the same channel. The communication range radius of all nodes is 35 m. We assume that the signal-to-noise ratio (SNR) is 25 dB, which is the ratio of transmission power to noise power on the receiver side. The simulation parameters are listed in Table 2. To collect offline RSS information, we conducted realistic experiments based on IEEE 802.15.4/Zigbee applications using a CC2530 system, as shown in Figure 5. The ZigBee nodes in our network are divided into two types: a coordinator and end devices. We set the coordinator to the sink node, which collects all sensing information in the network and sends information to the base station. An end device is an anchor node, target node, or malicious node.

Localization error is defined as the Euclidean distance between a true position and estimated position. This distance is calculated as follows:

$$\delta = \sqrt{(R_x - P_x)^2 + (R_y - P_y)^2}, \tag{16}$$

where δ is the location error, (R_x, R_y) are the real coordinates of an unknown node, and (P_x, P_y) are the estimated coordinates of an unknown node.

In Figure 6, the red curve represents the simulation results of the proposed algorithm (CSMR_SL), the blue curve represents the CS-based secure localization algorithm (CS_SL), and the green curve represents the CS-based source localization algorithm without security protection (CS_NSL). In this graph, the horizontal axis represents the number of target nodes. Fifteen target nodes were tested in the 100 × 100 m^2 observation area. The vertical axis represents the positioning error. These values represent the distances between the real positions and estimated positions. The positioning error for each number of target nodes represents the average of 16 experimental trials. From this graph, one can see that with an increase in the number of target nodes, the mutual interference between nodes becomes stronger, which leads to the upward trend in positioning error. With a uniform increase in the number of target nodes, the increase in localization error is very small. Compared to the other two methods, the proposed algorithm has clear advantages. The maximum error values for the proposed method are 41.8% and 61.3% lower than those for the CS_SL and CS_NSL methods, respectively.

TABLE 2 Simulation parameters

Parameter	Value
Observation area (m^2)	100 × 100
SNR (dB)	25
Communication radius (m)	35
Anchor numbers	17
Target numbers	Random
Anchor arrangement	Equilateral triangle
Length of equilateral triangle sides (m)	10–18
Target position	Random

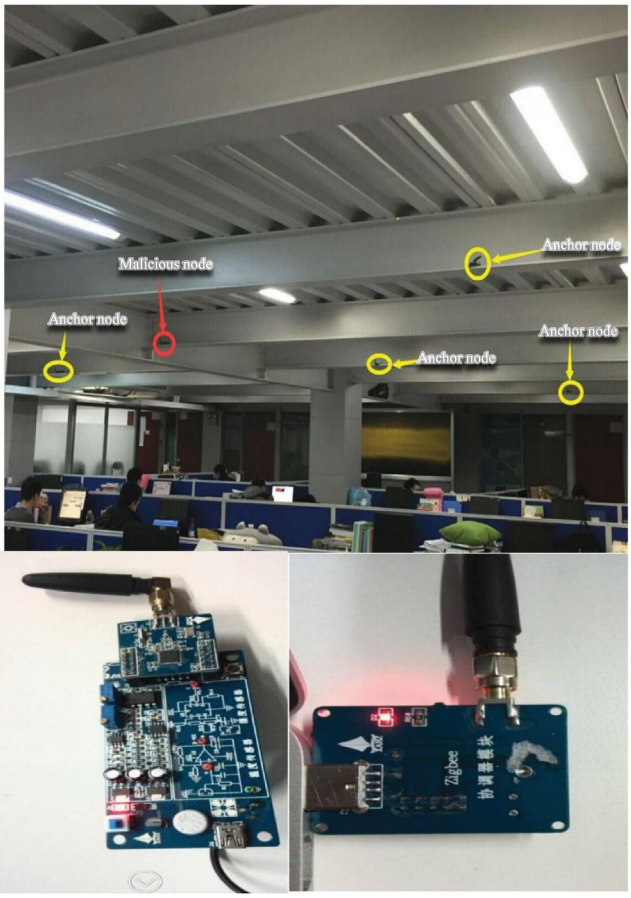

FIGURE 5 Realistic scenario with a vibration sensor and CC2530 system

Figure 7 presents experimental results for the three algorithms in terms of their ability to resist attacks. In this study, we tested i malicious anchor nodes with values of i ranging from one to six. Because malicious anchor nodes may have uneven distributions, they can lead to large deviations in positioning results. In this experiment, the target node was fixed and the i malicious anchor nodes were randomly placed in the observation area. Each data point represents the average value from 10 experimental trials. As shown in Figure 7,

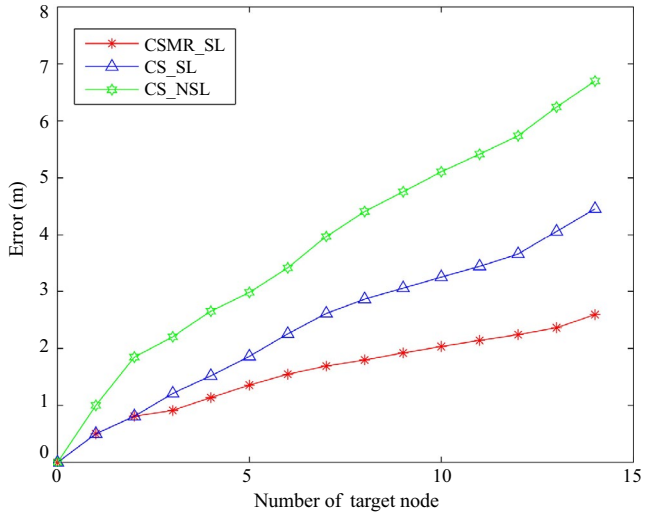

FIGURE 6 Localization error versus number of targets

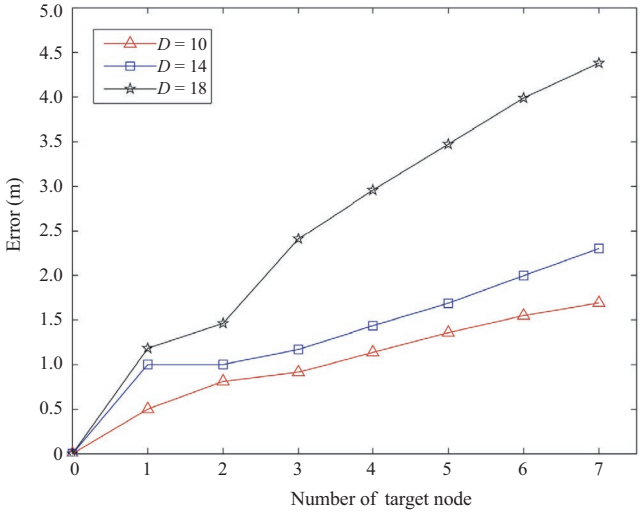

FIGURE 8 Localization error versus length of equilateral triangle sides

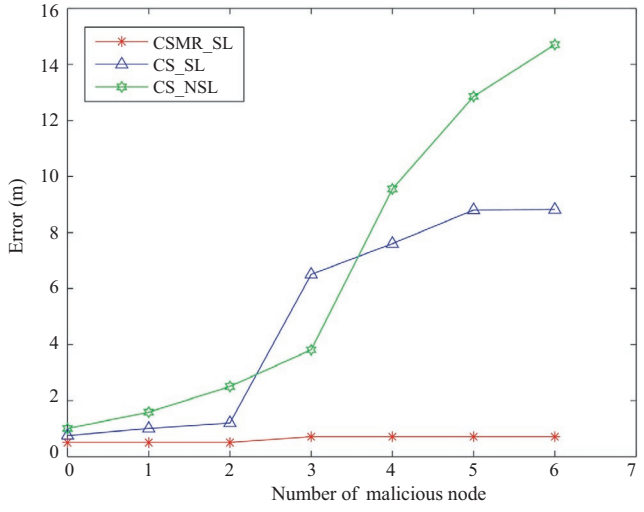

FIGURE 7 Localization error versus number of malicious nodes

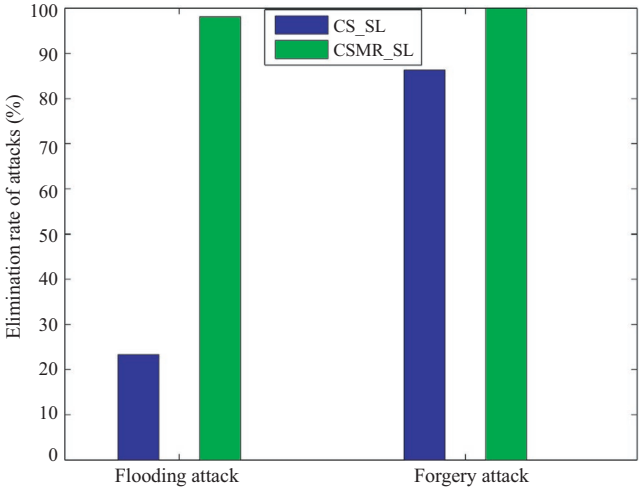

FIGURE 9 Elimination rates for forgery attacks and replay attacks

when the number of malicious nodes is small (one or two), the location errors for all three algorithms are similar. When the number of malicious nodes is three, the locating error of CS_NSL is greater than that of CS_SL. The main goal of CS is to reconstruct sparse signals with a relatively low sampling frequency. A localization algorithm based on CS theory can reduce redundant information by extracting high-quality RSS readings. Because malicious nodes are deployed randomly in the network, some nodes may be placed at boundary locations, causing the signal processing and security mechanism of CS to produce misjudgments during location analysis, resulting in location errors. Overall, the location accuracy of CS_SL is lower than that of CS_NSL. However, CSMR_SL algorithm proposed in this paper uses multiregion technology to reduce the location area and optimize location information in the network. Therefore, regardless of malicious node positioning, better location accuracy can be obtained by CSMR_SL. When the number of malicious nodes is greater than three, the errors of CS_NSL and CS_SL show sharp increases. With five or more malicious anchor nodes, the growth in CS_SL positioning error is relatively flat, but the error values are still very large. The location error of the CSMR_SL algorithm does not appear to increase dramatically and shows much lower values compared to the errors of the other two methods.

The relationship between the length of the equilateral triangle sides (D) and the location error for different numbers of target nodes is presented in Figure 8. It should be noted that with an increase in the number of target nodes, the location error of the CSMR_SL algorithm increases dramatically when the length of the equilateral triangle sides is 18 m. This indicates that the length of the equilateral triangle sides and the number of target nodes are proportional to the location error. As the length of the equilateral triangle sides and the

number of target nodes increase, indoor location becomes increasingly inaccurate.

As shown in Figure 9, we tested CSMR_SL and CS_SL under forgery attacks and replay attacks. The figure clearly shows that CSMR_SL offers strong defense against forgery attacks and replay attacks. In contrast, CS_SL is only effective at defending against forgery attacks. Unlike CS_SL, where secret keys are shared across the entire network, the preinstalled secret keys in CSMR_SL are shared according to region deployment, meaning different regions have different secret keys. A state analysis revealed that the multiregional secure mechanism is more sensitive in terms of detecting and defending against attacks.

5 | CONCLUSION

This study investigated the problem of sparse target positioning in indoor WSNs containing malicious nodes. We proposed the CSMR_SL algorithm to reduce the effects of hostile indoor environments and provide more precise room-level localization by overlapping sensing anchor communication regions to update a redundant dictionary dynamically. Additionally, weighted grids are incorporated to locate unknown nodes prior to executing the CS algorithm. We proposed a multiregion mechanism with transmission-range-based secret keys to prevent attacks by malicious nodes. Simulations of target localization demonstrated that the proposed algorithm can achieve higher accuracy and better performance in hostile indoor environments compared to algorithms without security protection.

Mobile target tracking and navigation using WSNs have attracted the attention of many researchers. WSNs are cheap and can be easily deployed. Therefore, they are widely used in location-based services. To provide such services, there is significant interest in developing real-time and accurate indoor tracking systems. Although indoor environments are often complex and hostile, the proposed CSMR_SL algorithm reduces the complexity of multitarget localization. Therefore, future work will focus on improving the tracking accuracy and security of mobile targets with lower energy consumption.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (61672220) and Key Technology Research and Development Plan of Hunan (2017GK2030).

ORCID

Xiangju Yao https://orcid.org/0000-0002-9187-2378

REFERENCES

1. Y. Li et al., *Qiloc: A qi wireless charging based system for robust user-initiated indoor location services*, in IEEE Int. Conf. Sens., Commun., Netw., Seattle, WA, USA, June 2014, pp. 184–185.
2. L. Tao, Z. Li, and L. Wu, *Outlet: outsourcing wearable computing to the ambient mobile computing edge*, IEEE Access **6** (2018), 18408–18419.
3. L. Yin, J. Luo, and H. Luo, *Tasks scheduling and resource allocation in fog computing based on containers for smart manufacturing*, IEEE Trans. Ind. Inform. **14** (2018), 4712–4721.
4. R. Garg, A. Varna, and M. Wu, *An efficient gradient descent approach to secure localization in resource constrained wireless sensor networks*, IEEE Trans. Inform. Forensics Security **7** (2012), 717–730.
5. S. Go and J. Chong, *Improved TOA-based localization method with BS selection scheme for wireless sensor networks*, ETRI J. **37** (2015), 707–716.
6. D. E. Chaitanya and G. S. Rao, *Unknown radio source localization based on a modified closed form solution using TDOA measurement technique*, Procedia Comput. Sci. **87** (2016), 184–189.
7. N. A. Khanbashi et al., *Measurements and analysis of fingerprinting structures for WLAN localization systems*, ETRI J. **38** (2016), 634–644.
8. S. Sorour et al., *Joint indoor localization and radio map construction with limited deployment load*, IEEE Trans. Mobile Comput. **14** (2015), 1031–1043.
9. C. Feng, S. Valaee, and Z. Tan, *Multiple target localization using compressive sensing*, in Global Telecommun. Conf., Honolulu, HI, USA, 2009, pp. 1–6.
10. C. Feng et al., *Compressive sensing based positioning using RSS of WLAN access points*, in IEEE INFOCOM, San Diego, CA, USA, Mar. 2010, pp. 1–9.
11. E. J. Candès and M. B. Wakin, *An introduction to compressive sensing*, Signal Process. Mag. **25** (2008), 21–30.
12. L. Liu et al., *Adaptive source location estimation based on compressed sensing in wireless sensor networks*, Int. J. Distrib. Sens. Netw. **8** (2012), 141–149.
13. Y. Mo et al., *A spatial division clustering method and low dimensional feature extraction technique based indoor positioning system*, Sens. **14** (2014), 1850–1876.
14. M. Jadliwala et al., *Secure distance-based localization in the presence of cheating beacon nodes*, IEEE Trans. Mobile Comput. **9** (2010), 810–823.
15. J. Luo et al., *Secure indoor localization based on extracting trusted fingerprint*, Sens. **18** (2018), 469–492.
16. Y. Chen et al., *Indoor localization using FM signals*, IEEE Trans. Mobile Comput. **12** (2013), 1502–1517.
17. T. Higuchi et al., *Mobile node localization focusing on stop-and-go behavior of indoor pedestrians*, IEEE Trans. Mobile Comput. **13** (2014), 1564–1578.
18. F. Anjum, S. Pandey, and P. Agrawal, *Secure localization in sensor networks using transmission range variation*, in IEEE Int. Conf. Mobile Adhoc Sens. Syst., Washington, DC, USA, Nov. 2005, pp. 9–203.
19. B. Sklar, *Digital communications*, Prentice-Hall, Upper Saddle River, NJ, USA, 2001, p. 190.
20. R. Baraniuk et al., *A simple proof of the restricted isometry property for random matrices*, Constructive Approximation **28** (2008), 253–263.

21. J. Romberg, *Imaging via compressive sampling*, IEEE Signal Process. Mag. **25** (2008), 14–20.
22. B. Zhang et al., *Sparse target counting and localization in sensor networks based on compressive sensing*, in IEEE INFOCOM, Shanghai, China, Apr. 2011, pp. 2255–2263.
23. E. Candè and J. Romberg, *Sparsity and incoherence in compressive sampling*, Inverse Prob. **23** (2006), 969–985.
24. D. Milioris et al., *Low-dimensional signal-strength fingerprint-based positioning in wireless LANs*, Ad hoc Netw. **12** (2014), 100–114.
25. C. Feng, S. Valaee, and Z. Tan, *Multiple target localization using compressive sensing*, in Global Telecommun. Conf., Honolulu, HI, USA, 2009, pp. 1–6.
26. J. Jiang et al., *Secure localization in wireless sensor networks: A survey*, J. Commun. **6** (2011), 460–470.

AUTHOR BIOGRAPHIES

Chang Liu received her BS and MS degrees from the Changchun University of Science and Technology, Changchun, Jilin, China in 2009 and 2014, respectively. She is currently pursuing a PhD at the College of Computer Science and Electronic Engineering, Hunan University, Changsha, Hunan, China. Her research interests include wireless networks.

Xiangju Yao received her MS degree from the College of Computer Science and Electronic Engineering, Hunan University, Changsha, Hunan, China in 2017. Her research interests include wireless networks.

Juan Luo received her BS degree from the National University of Defense Technology, Hunan, China in 1997, and her MS degree and PhD in Communication and Information Systems from Wuhan University, Hubei, China in 2000 and 2005, respectively. She is currently a professor and doctoral supervisor at the College of Computer Science and Electronic Engineering, Hunan University, Changsha, Hunan, China. From 2008 to 2009, she was a visiting scholar at the University of California at Irvine. She has published more than 60 papers. Her research is focused on wireless networks, cloud computing, and the internet of things. She is a member of IEEE, SIGCOM, and ACM, and a senior member of CCF.