An Attack-Resistant Weighted Least Squares Localization Algorithm Based on RSSI

Yitong Liu[1], Jun Peng[2], Xingcheng Liu[2,3(✉)], Yi Xie[1], and Zhao Tang[2]

[1] School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China
{isslxc,xieyi5}@mail.sysu.edu.cn
[2] School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China
[3] School of Information Science, Guangzhou Xinhua University, Guangzhou, China

Abstract. As an important part of the Internet of things (IoT), wireless sensor networks (WSNs) have been applied in many fields. Most applications require accurate location information, hence node localization is one of the important issues in WSNs. It is very important to ensure the security of localization when WSNs are under attack. A new attack-resistant weighted least squares (ARWLS) algorithm based on RSSI was proposed in the paper. The algorithm is oriented to the problem solution for the situation that the attacker influences the system by tampering with the transmitting power in the localization mechanism. The proposed algorithm can be used in the attack scenarios. Simulations results show that, compared with other algorithms, the proposed algorithm has merits in localization accuracy and robustness to resisting the tampering activities of attackers.

Keywords: Wireless sensor networks · Malicious nodes · Secure localization · Sequential probability ratio test · Weighted least squares

1 Introduction

With the continuous development of related technologies, the Internet of Things (IoT) is playing an increasingly important role in people's daily lives in recent years [1]. As an important part of the IoT, a wireless sensor network (WSN) has a high research value [2]. It has been used in many fields, such as national defense, environmental monitoring, medical health, mechanical fault diagnosis and so on [3,4]. Most applications require accurate location information, and the

The work was supported by the Joint Key Program of the National Natural Science Foundation of China and Guangdong Province of China(Grant No. U2001204), by the National Natural Science Foundation of China (Grant Nos. 61873290, 61972431 and 61572534), and by the Science and Technology Program of Guangzhou, China (Grant No. 202002030470).

S. Paiva et al. (Eds.): SmartCity360° 2020, LNICST 372, pp. 476–494, 2021.
https://doi.org/10.1007/978-3-030-76063-2_32

information collected by the sensor nodes cannot be processed correctly without location information of themselves. Equipping the node with a global positioning system(GPS) device or deploying the node in a predetermined location is the direct mean to obtain the location information. However, the resources of WSNs are limited, the power consumption of equipping each node with GPS devices is too high. In addition, under nonline-of-sight (NLOS) conditions, the error caused by GPS becomes larger [5,6]. What's more, WSNs are often deployed in harsh and dynamically changing environments, and it is not practical to deploy each sensor node in a specific location. In practice, only part of the node's location information is known, which is called the anchor node. It is a critical technique in WSN to estimate the location of the target node by using prior information. Currently, the localization algorithms can be divided into two categories: the range-based [7–10] and the range-free [11–13] localization algorithms. For the range-based algorithm, the distance or angle values are directly measured and computed through the physical ranging technologies. For the range-free localization algorithms, the location estimate of the target node relies on the connectivity of the whole network topology.

1.1 Related Works

In recent decades, various security strategies have been proposed to address the node location in malicious attack environments. The strategies for resisting attacks vary according to the application scenarios. The attack-resistant localization strategies proposed in literatures can be divided into three categories: malicious anchor node detection algorithms [14–16], robust localization algorithms [17–20] and location verification strategies [21–23].

In the process of location estimation of the target node, the information provided by the anchor node is required, so the reliability of the anchor node largely determines the reliability of the location result. The method of malicious anchor node detection is designed to filter the information provided by the malicious anchor node by analyzing the network model and the behavior characteristics of the malicious anchor node, and then to estimate the location of the target node with the anchor node with high reliability. To this end, a malicious node detection algorithm based on clustering and consistency evaluation (MNDC) [14] was proposed. The algorithm takes advantage of the consistency between the distance measurements of ToA and RSSI. The differences between the distance measurements of ToA and RSSI are taken as the basis of the detection. With the aid of the designed scheme, the anchor nodes under malicious attack when using ToA are eliminated. However, the limitation of this algorithm lies in the need to ensure that RSSI measurements are always protected from attack. Grag. et al. proposed an efficient Gradient Descent (GD) approach [15]. The approach is divided into two stages. In the first stage, the least squares solution is searched by gradient descent method. When the sum of the gradients of anchor nodes exceeds the preset threshold, the selection pruning stage is entered. 50% anchor nodes with larger gradient are regarded as malicious, while the remaining anchor nodes are adopted to continue the iteration.

The ability of the algorithm to tolerate attacks can be improved by improving the robustness of the distance measurement stage and the position calculation stage. Improving the robustness of distance measurement is usually achieved through time or space constraints. Distance-Bounding Protocol [17] was proposed, which can prevent distance shortening attacks caused by early response of nodes. However, there is a big limitation in the distance boundary protocol. The protocol is based on the time of arrival of the RF signal, leading to an extremely high requirement on the accuracy of the recorded time. The nanosecond error in the time is represented as a difference of tens of centimeters in the distance. At present, most localization algorithms adopt Least Squares (LS) in the final estimation of position calculation stage. The essence of the LS algorithm is to find the coordinates of the target nodes corresponding to the minimum sum of residual of all anchor nodes. LS is sensitive to outliers. Once an anchor node is under attack, it may cause a relatively large deviation in the estimated location. To avoid that, Li et al. [18] divided the anchor nodes into several subsets, and in each subset Least Median Squares (LMS) is used to get the corresponding location candidate values. The candidates with the minimum residual value is regarded as the location estimation. In addition, they proposed an adaptive LS and LMS positioning mechanism, among which the LS has the computational advantage in the absence of attacks. To solve the secure localization problem, a weighted Least Squares (WLS) algorithm [19] was proposed based on RSSI. To go further, an attack-resistant localization algorithm [20], was proposed. The consistency of multiple anchor nodes can be used in the same network to achieve secure localization. Meanwhile, a voting-based localization algorithm [20] was presented to treat the problem. According to the coverage of each anchor node, the area to be detected can be partitioned into grids. The grids are judged and rated with a vote. Finally, the center of the grid with the highest number of votes is selected as the final location estimation of the target node.

Location verification system is also used to improve the security of the networks. The location anomaly detection (LAD) algorithm [23] determines whether the node is malicious by judging whether the error between the estimation and the real location is within a certain threshold. If the difference value does not exceed the threshold, then a decision is made that the node is benign. However, the prior information of node distribution is required. Therefore, LAD does not work in many scenarios, the final solution of which is to be resolved.

1.2 Contributions

In order to solve the security problem of range-based localization mechanism under attack, we proposed an attack-resistant weighted least squares (ARWLS) localization algorithm. ARWLS algorithm operation consists of an anchor node screening stage and a final location calculation stage. In the proposed algorithm, the main contributions can be summarized as follows:

- A scheme is proposed to identify the malicious anchor nodes.
- The quasi-Newton [24] method is used to determine the reference anchor nodes, and the initial location estimate, which is the basis of the followed detection, but not the whole detection itself.
- The estimate of the initial location is used to approximate the real location of the target node. Then the power difference information is used as the sample of sequential probability ratio test (SPRT).

Compared with other secure localization algorithms, the proposed algorithm guarantees the security of location estimation from the two aspects: eliminating or decreasing the malicious anchor nodes and improving the robustness of the location calculation. In addition, the proposed algorithm has no requirement of prior information or additional hardware facilities.

The rest of the paper is organized as follows. The network model and the localization problem formulations under malicious attack are demonstrated in Sect. 2. The idea and the specific process of the proposed algorithm are described in detail in Sect. 3. Then the simulation results and corresponding analysis are presented in Sect. 4. Finally, Sect. 5 summarize the work of the paper.

2 Network Model and Problem Formulations

2.1 Network Model

We consider a two-dimensional localization system in which all anchor nodes in the region are randomly distributed. It is assumed that the system satisfies the following conditions.

1) The network is stable and the locations of all the sensor nodes are not changed after deployed;
2) All anchor nodes are within the communication range of the target node, so the target node can receive signals from all anchor nodes.

2.2 Problem Formulations

It is assumed that there are N anchor nodes with known locations in the network. The location of the i-th anchor is denoted by $\mathbf{A}_i = [x_i, y_i]^T, i = 1, 2 \cdots N$. Among the N anchor nodes, M of them are malicious, which will send false messages to interfere with the localization process. Assume that there is a target node to be located in the system, and its real position is denoted by $\mathbf{T} = [x_t, y_t]^T$. Assuming that the anchor nodes send messages at a fixed power level, the target node is located according to the received signal strength indication and the location of the anchor nodes. The true distance between the anchor node and the target node is represented with $d = [d_1, d_2, \cdots d_N]^T$, where $d_i = \|\mathbf{A}_i - \mathbf{T}\|$. For RSSI ranging, the distance between the anchor nodes and the target node can be estimated through the path loss in the transmission process, which is related to

the specific transmission model. In this paper, the log-distance model is adopted, and the received power can be modeled as Eq. (1) [25]:

$$p^r = p_0^t - 10a\log_{10}(d). \tag{1}$$

where p_0^t(dBm) represents the transmitting power of the anchor node at a reference distance, p^r represents the received power of the anchor nodes, a denotes the path loss exponent, and d is the distance between the anchor node and the target node. Equation (1) describes the received power under ideal conditions. The target node receives P packets from the i-th anchor node, and the corresponding received power can be denoted by $\mathbf{p}_i^r = [p_{i1}^r, p_{i2}^r, \ldots, p_{iP}^r]^T, i = 1, 2, \ldots, N$. Given the transmitting power p_0^t and received power p_{ij}^r, the corresponding distance measurement ed_{ij} can be estimated, as shown in Eq. (2):

$$ed_{ij} = 10^{\frac{p_0^t - p_{ij}^r}{10a}}, i = 1, 2, \ldots, N, j = 1, 2, \ldots, P. \tag{2}$$

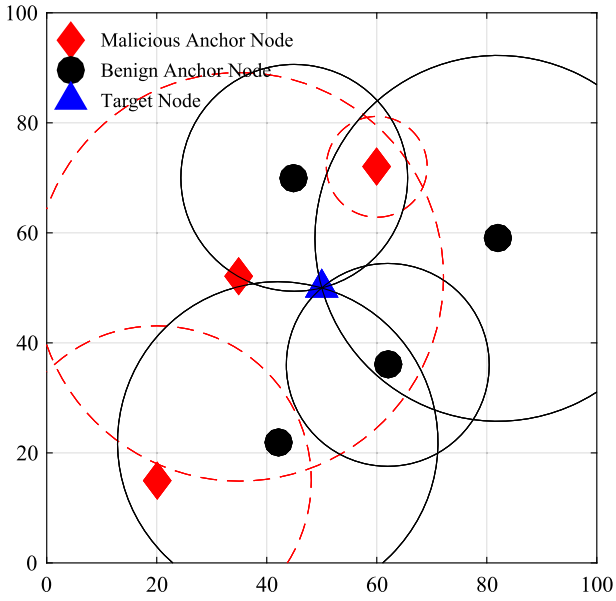
According to Eq. (2), transmitting power transmitting power p_0^t, received power p_{ij}^r and the path loss exponent a all have an impact on the distance measurement. Based on Eq. (2), it can be inferred that the relationship between the received signal power and the measured distance is nonlinear.

Noncoordinated Attacks. In a noncoordinated attack environment, the attacker acts alone on each captured anchor node, tampering with its transmitting power, and the target node is unaware whether the transmitting power of the anchor node has been changed. In this paper, we model this attack scenario by reporting the received signal power of the target node. Under noncoordinated attack, the received signal power at the target node can be defined as Eq. (3):

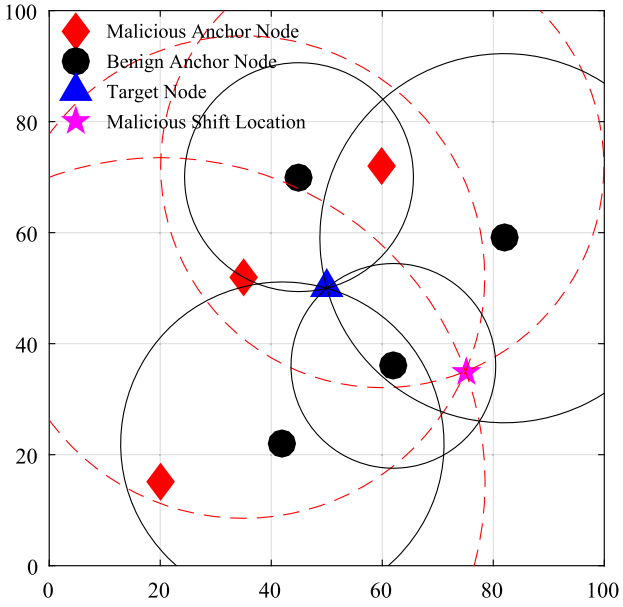
$$(p_i^r)^{(nc)} = \begin{cases} p_0^t - 10a\log_{10}(d_i) + n_i, & \text{if } M_i = 0 \\ p_{m_i}^{nc} - 10a\log_{10}(d_i) + n_i, & \text{if } M_i = 1 \end{cases} \tag{3}$$

where $M_i = 0$ denotes the case that the i-th anchor node is benign, and $M_i = 1$ denotes the i-th anchor node is malicious. p_0^t represents the predefined transmitting power, and all the benign anchor nodes broadcast packets. A zero-mean Gaussian random variable n_i is used to model measurement noise, and the variance of n_i is σ^2, i.e. $n_i \sim \mathcal{N}(0, \sigma^2)$. n_i is influenced by environmental factors. The transmitting power of the malicious anchor node is tampered to p_m^t by the attacker. $p_{m_i}^{nc}$ is defined as $p_{m_i}^{nc} = p_0^t + \kappa$, where κ is a Gaussian random variable with a mean of zero and variance of σ_{att}^2, i.e. $\kappa \sim \mathcal{N}(0, \sigma_{att}^2)$. The number of malicious anchor nodes, m, and the standard deviation of attack term, σ_{att}, will affect the degree of attack to the whole network.

Coordinated Attacks. In a coordinated attack environment, the attacker may specify a position, which may be randomly selected, or a position that is determined to be favorable to the attacker. The specified position can be denoted by $\mathbf{T}_{mal} = [x_m, y_m]^T$. Multiple malicious anchor nodes communicate with each



(a) Noncoordinated Attacks



(b) Coordinated Attacks

Fig. 1. Attacks caused by malicious anchors in a WSN with 7 anchor nodes and 3 of them are malicious, ignoring the measurement noise n.

other to reach cooperation and attempt to locate the target node to $\mathbf{T}_{mal}$. In this scenario, the received power $(p_i^r)^c$ can be modeled as:

$$(p_i^r)^{(c)} = \begin{cases} p_0^t - 10a\log_{10}(d_i) + n_i, & \text{if } M_i = 0 \\ p_{m_i}^c - 10a\log_{10}(d_i) + n_i, & \text{if } M_i = 1 \end{cases} \tag{4}$$

where $p_{m_i}^c$ is defined by the distance between $\mathbf{T}_{mal}$ and $\mathbf{T}$. We define $d_a = \|\mathbf{T}_{mal} - \mathbf{T}\|$. Therefore, $p_{m_i}^c$ can be defined as $p_{m_i}^c = p_0^t - 10a\log_{10}(\frac{\|\mathbf{T}_{mal} - \mathbf{A}_i\|}{d_i})$. In fact, the goal of the coordinated attack is to change the distance measurements from the malicious anchor nodes to the target node $\mathbf{T}$ into the distance to the specified position $\mathbf{T}_{mal}$ by tampering with the transmitting power. The degree of coordinated attack can be represented by the value of d_a.

Figure 1 shows the models of the uncoordinated attack and coordinated attack, where the measurement noise is 0. Suppose that there are seven anchor nodes in the system, three of which are malicious. In Fig. 1(a), the three malicious anchor nodes act alone. In Fig. 1(b), the three malicious anchor nodes attempt to shift the target node to the specified location.

3 Proposed Algorithm for Secure Localization

Aiming at the situation that the attacker tampers with the transmitting power of anchor node based on RSSI, we proposed an attack-resistant weighted least squares (ARWLS) to achieve secure localization.

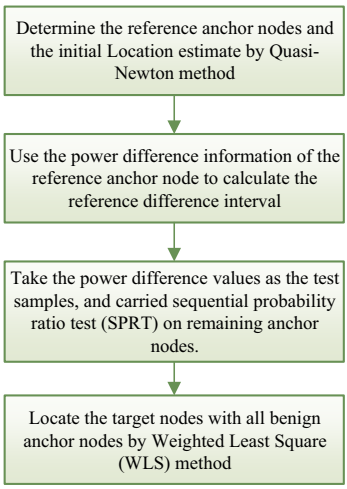


Fig. 2. Main steps of ARWLS algorithm

The main steps of ARWLS algorithm are shown in Fig. 2. The algorithm can be divided into two stages. The first stage is the anchor node screening. At this stage, we remove the detected malicious anchor nodes. This stage consists of three steps. First, assuming that there is no malicious anchor node, the residual sum of all nodes is taken as the objective function and the quasi-Newton method is adopted to obtain the initial location estimate of the target node. According to the characteristics of the initial location estimate, which is close to the real location of the target node, the anchor nodes with smaller gradients are regarded as the reference anchor nodes. Then, based on the distance measurements and the initial location estimate, the approximate value of the difference between the real transmitting power and the transmitting power in the non-attacking state can be calculated. According to the power difference information of the reference anchor nodes, a reference error interval can be calculated, and the judgment of the remaining anchor nodes is regarded as a hypothesis testing problem [26,27]. The power difference values are taken as the test samples, with which the method of sequential probability ratio test (SPRT) [28] is adopted to detect the remaining anchor nodes. In the second stage, the benign anchor nodes are used for positioning. The mean value and variance of multiple measurements are used to calculate the corresponding weight of each benign anchor node, and the weighted least square method is used for the final location estimate of the target node.

In the first stage, the malicious anchor nodes are eliminated to prevent the false information from being used in the localization of the target nodes and improve the capability of the algorithm in resisting attacks. In the second stage, the robustness of location calculation is improved. Even if some malicious anchor nodes are not removed in the first stage and then allowed to participate in the localization process, the final location estimate will not be greatly affected. Therefore, the algorithm improves the security of location from the above two aspects. Next, we will elaborate on each step in detail. Figure 3 is the flow chart of the whole ARWLS algorithm.

3.1 Anchor Nodes Screening Stage

In this stage, the received power information corresponding to all anchor nodes is collected. Therefore, the average measured distance corresponding to each anchor node $\overline{ed}_i, i = 1, 2, \ldots, N$ can be calculated. The satge consists of the following three steps: the reference anchors determined, the reference interval computed, and the SPRT operation performed.

Determine the Reference Anchor Nodes and the Initial Location Estimate. Firstly, the BFGS which is one of quasi-Newton methods is used to obtain the LS solution when all anchor nodes participate in the localization of the target nodes. The BFGS algorithm has a superlinear convergence speed, which can converge to the LS solution faster than the gradient descent (GD) algorithm. Assuming there is no malicious anchor node, the BFGS algorithm is then used

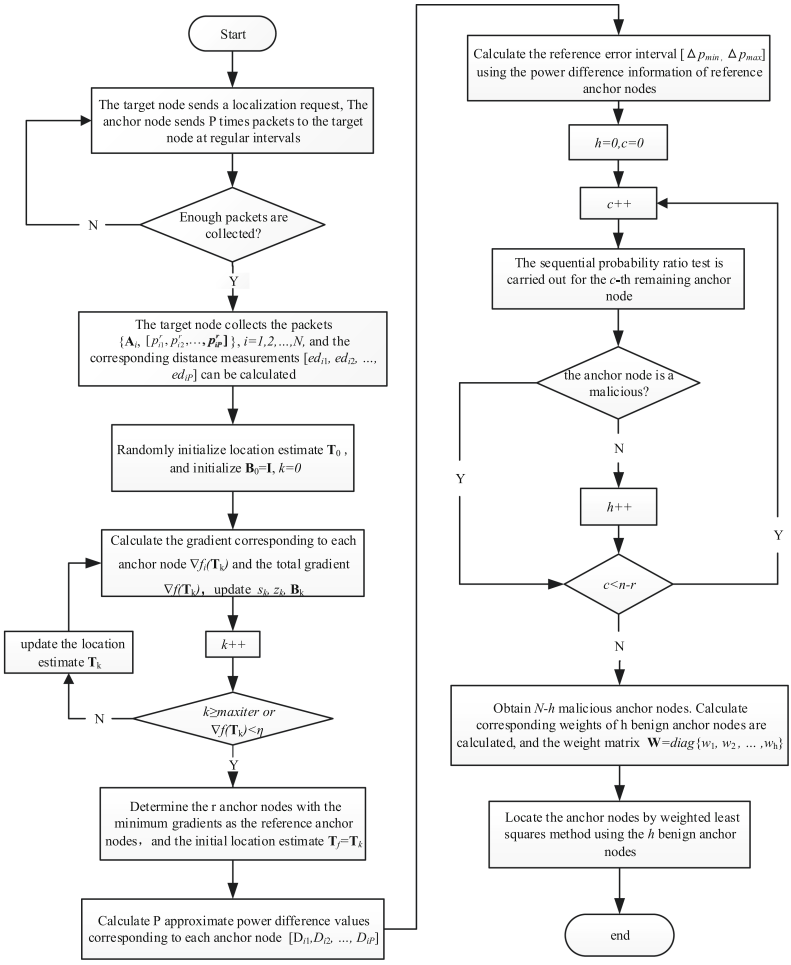


Fig. 3. Flowchart of ARWLS algorithm

to iteratively search the LS solutions. The residual sum of all anchor nodes is taken as the objective function, as shown in Eq. 5:

$$\hat{\mathbf{T}} = \arg\min_{\mathbf{T}} \sum_{i=1}^{N} (\|\mathbf{A}_i - \mathbf{T}\| - \overline{ed}_i)^2 = \arg\min_{\mathbf{T}} f(\mathbf{T}). \tag{5}$$

According to the BFGS algorithm, we have,

$$\mathbf{B}_{k+1} \left[\nabla f(\mathbf{T}_{k+1}) - \nabla f(\mathbf{T}_k)\right] \approx \mathbf{T}_{k+1} - \mathbf{T}_k, \tag{6}$$

where k denotes the number of iterations, $\mathbf{B}_k$ is approximately positive definite matrix of the inverse of the Hessian matrix $\mathbf{H}_k^{-1}$, $\nabla f(\mathbf{T}_k)$ represents the total gradient corresponding to the location of k-th iteration.

According to the relevant properties of the BFGS algorithm, the update formula of the location estimate $\mathbf{T}_k$ is shown in Eq. (7)

$$\mathbf{T}_{k+1} = \mathbf{T}_k - \lambda \nabla f(\mathbf{T}_{k+1}), \tag{7}$$

where λ denotes the step size for searching.

Calculate the Power Difference Reference Interval. Since the initial location estimate $\mathbf{T}_f$ is close to the real position of the target node, it can be used to approximate the real location of the target node $\mathbf{T}$. Given the received power p_{ij}^r and the approximate location of the target node $\mathbf{T}_f$, the j-th transmitting power of the i-th anchor can be approximately inferred with Eq. (8) if the measurement noise is ignored,

$$p_{ij}^t \approx 10a \log_{10}(\|\mathbf{A}_i - \mathbf{T}_f\|) + p_{ij}^r. \tag{8}$$

All anchor nodes send P packets at regular intervals. Then, each anchor node gets the corresponding P distance measurements. Given the distance measurements ed_{ij} and the predefined transmitting power p_0^r, the j-th transmitting power of the i-th anchor can be approximately inferred with Eq. (9) if the measurement noise is ignored,

$$\hat{p}_{ij} \approx 10a \log_{10}(ed_{ij}) + p_{ij}^r. \tag{9}$$

Therefore, the power difference value can be calculated as Eq. (13):

$$\begin{aligned} D_{ij} &\approx \hat{p}_{ij} - p_{ij}^t \\ &= 10a \log_{10}\left(\frac{ed_{ij}}{\|\mathbf{A}_i - \mathbf{T}_f\|}\right). \end{aligned} \tag{10}$$

For benign anchor nodes, the difference is only affected by the error of the LS solution and measurement noise in the first stage. Compared with the error introduced by the malicious anchor nodes, this difference is very small. The set of reference anchor nodes is a subset of all anchor nodes. Based on the premise that the reference anchor nodes are benign, the difference information of the reference anchor node is taken as the basis for establishing the detection model. The power difference values are taken as the samples. Suppose a reference anchor node is an individual, then for the i-th individual, the mean value $\overline{D}_i$ and the variance s_i^2 of the individual can be calculated as follows:

$$\overline{D}_i = \frac{\sum_{j=1}^{P} D_{ij}}{P}, \tag{11}$$

$$s_i^2 = \frac{\sum_{j=1}^{P} (D_{ij} - \overline{D}_i)^2}{P-1}. \tag{12}$$

Averaging the mean values of the different individuals, the mean value of total samples can be calculated as Eq. (16):

$$\overline{D} = \sum_{i=1}^{r} \frac{D_i}{r}. \tag{13}$$

Next, the differences between individuals of each reference anchor node are analyzed, and the variation s_e^2 of the mean value $\overline{D}_i$ between individuals can be expressed as Eq. (14):

$$s_e^2 = \sum_{i=1}^{r} \frac{(\overline{D}_i - \overline{D})^2}{r-1}. \tag{14}$$

The variance between individuals s_a^2 is given in Eq. (15):

$$s_a^2 = \sum_{i=1}^{r} \frac{P-1}{N-r} s_i^2, \tag{15}$$

where $N = P \cdot r$ is the total number of samples. Based on the analysis of the s_i^2 and s_a^2, the population distribution of the sample can be described, and the population variance of the sample can be expressed as Eq. (16):

$$s_t^2 = s_e^2 + (1 - \frac{1}{m_h}) s_a^2, \tag{16}$$

where m_h is the harmonized mean of the numer of the measured times. Since each anchor node sends P packets at regular intervals, $m_h = \frac{r}{\frac{1}{P} \cdot r} = P$. Based on the above analysis, the reference error interval of the power difference between reference anchor nodes can be calculated as Eq. (17) and Eq. (18) when the significance level is ϵ:

$$D_{min} = \overline{D} - (z_{1-\frac{\epsilon}{2}}) \times \sqrt{s_t^2}, \tag{17}$$

$$D_{max} = \overline{D} + (z_{1-\frac{\epsilon}{2}}) \times \sqrt{s_t^2}, \tag{18}$$

where $z_{1-\frac{\epsilon}{2}}$ is the upper quartile of $1 - \frac{\epsilon}{2}$ of the standard normal distribution.

Sequential Probability Ratio Test (SPRT). According to the reference difference interval obtained in the previous step, a Bernoulli random variable Z_{ij} can be established for the difference information D_{ij} of the remaining anchor nodes, which is given in Eq. (19),

$$Z_{ij} = \begin{cases} 0, D_{min} < D_{ij} < D_{max} \\ 1, \quad others \end{cases}. \tag{19}$$

Define the probability of D_{ij} exceeding the reference difference interval as p, i.e. $P(Z_{ij} = 1) = p$. So the probability that D_ij is within the reference difference interval is $1 - p$, i.e. $P(Z_{ij} = 0) = 1 - p$. Two hypotheses H_0 and H_1 can be established:

- H_0: the detected anchor node is benign, $p \leq p_0$,
- H_1: the detected anchor node is malicious, $p > p_1$,

where p_0 and p_1 are preset thresholds.

Through the above test, the malicious anchor nodes can be screened out, and the information provided by them will be eliminated, which is helpful to improve the localization accuracy.

3.2 Location Calculation Stage

For benign anchor nodes, there exists a nonlinear relationship between the received power and the distance measurements. For the anchor node which is far from the target node, the noise of the same size will cause more fluctuation in distance measurements. Therefore, it can be concluded that the anchor nodes close to the target node are more robust to noise [19]. For the malicious nodes, its transmitting power is affected by the extra attack item, so its fluctuation range of the distance measurements is larger.

Firstly, the selected benign anchor nodes are relabeled as $1, 2, \cdots, h$, and the corresponding coordinates are denoted as $(x_i, y_i), i = 1, \cdots, h$. We use the variance of the squares of distance measurements $\mathbf{ed}_i^2 = [ed_{i1}^2, ed_{i2}^2, \ldots, ed_{iP}^2]^T, i = 1, 2, \ldots, N$ as the standard to measure the reliability of anchor nodes. The variance $Var(ed_{ij})^2$ can be calculated as Eq. (20):

$$Var(ed_{ij}^2, \overline{ed_i^2}) = \frac{\sum_{j=1}^{P}(ed_{ij}^2 - \overline{ed_i^2})^2}{P-1}. \tag{20}$$

Each anchor node can be weighted by the variance of the square of the distance measurements $Var(ed_{ij})^2$, where the weight of the i-th anchor node can be calculated as Eq. (21)

$$w_i = \frac{1}{Var(ed_{ij}^2, \overline{ed_i^2})}. \tag{21}$$

The weight matrix can be defined as $\mathbf{W} = diag[w_1, w_2, \ldots, w_h]$. We can use a modified version of the least squares (LS) which is called weighted least squares (WLS) to make the final calculation. The final estimate of the location of the target node can be calculated by $\mathbf{WAt} = \mathbf{b}$, where the matrixes (vectors) are given in Eq. (22)

$$\mathbf{A} = \begin{bmatrix} -2x_1 & -2y_1 & 1 \\ -2x_2 & -2y_2 & 1 \\ \vdots & \vdots & \vdots \\ -2x_h & -2y_h & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} \overline{ed}_1 - x_1^2 - y_1^2 \\ \overline{ed}_2 - x_2^2 - y_2^2 \\ \vdots \\ \overline{ed}_h - x_h^2 - y_h^2 \end{bmatrix}. \tag{22}$$

4 Performance Evaluation

The proposed algorithm is experimently compared with two existing secure localization methods, namely, the Gradient Descent (GD) [15] and the Weighted Least Squares (WLS) [19] in this paper. The GD algorithm includes fix-step Gradient Descent (GD_f) algorithm and variable-step Gradient Descent (GD_v) algorithm, both of which have similar performance in localization accuracy. Here, we only show the performance of GD_f algorithm in this section. In WLS scheme, the prior information of the standard deviation of noise σ is required. In this section, the complexity of the algorithm is also discussed.

4.1 Experimental Platform and Parameters Setting

In this paper, all of the simulation experiments were performed on Mathworks MATLAB 2016a. The simulation environment is as follows: Intel Core I7-8700 CPU @3.20GHZ and 16GB RAM running with Windows 10 64-bit operating system. The parameters setting in the experiment is shown in Table 1 [19], except stated otherwise.

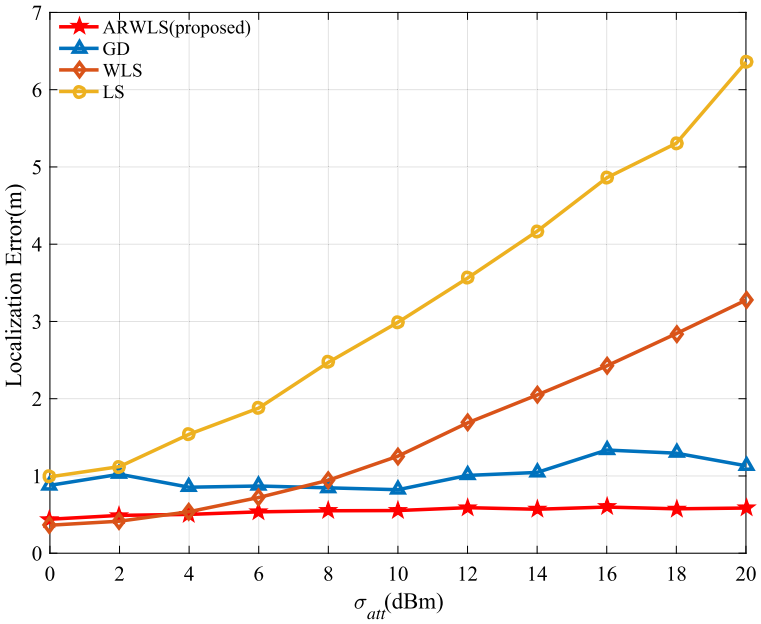
Table 1. Parameters setting

Symbols	Meanings	Values
N	Number of anchor nodes	30
M	Number of malicious anchor nodes	9
r	Number of reference anchor nodes	3
p_0^t	Predifined transmitting power	-10 dBm
σ	Std deviation of measurement noise	2 dBm
σ_{att}	Std deviation of attack	8 dBm
d_a	Distance between the malicious shift location and the target node	12 m
a	Path loss exponent	4
η	Gradient threshold	1.8 m
λ	Step size	0.5
P	Number of packages	20
p_0	Null hypothesis threshold	0.1
p_1	Alternative hypothesis threshold	0.9
α	False positive rate	0.01
β	False negative rate	0.01

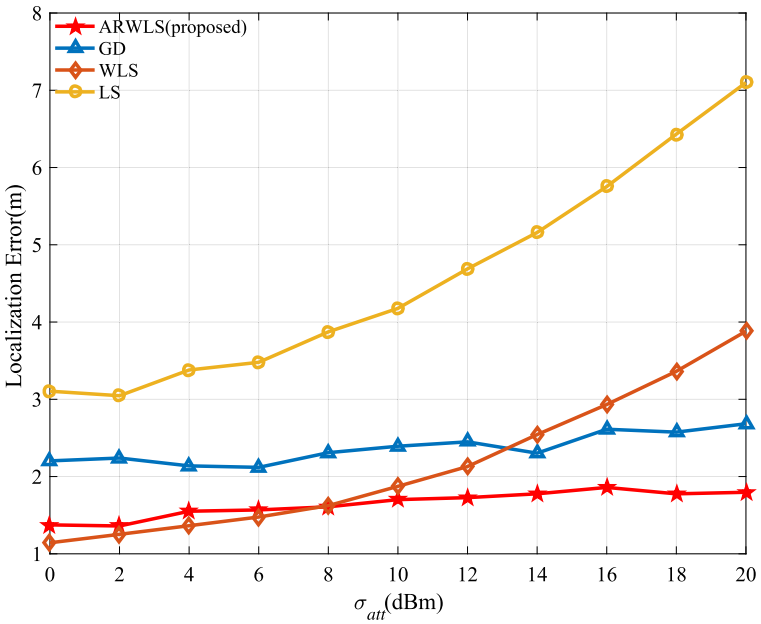
4.2 Simulation Results

The localization performance of the schemes under noncoordinated and coordinated attacks is demonstrated in this part.

Noncoordinated Attacks. The localization error of different algorithms varying with the std deviation of attack σ_{att} under different levels of noise σ is shown in Fig. 4. This figure shows the localization performance at $\sigma = 2$ dBm and $\sigma = 6$ dBm in Figs. 4(a) and 4(b), respectively. As is shown in the figure, the performance of the proposed ARWLS algorithm in positioning error is more stable than any other algorithms. Even if σ_{att} increases, the proposed algorithm can still ensure relatively high localization accuracy. This is because the proposed ARWLS algorithm eliminates the malicious anchor nodes in the screening stage



(a) Localization error with varied deviation of attacks at measurement noise σ=2dBm



(b) Localization error with varied deviation of attacks at measurement noise σ=6dBm

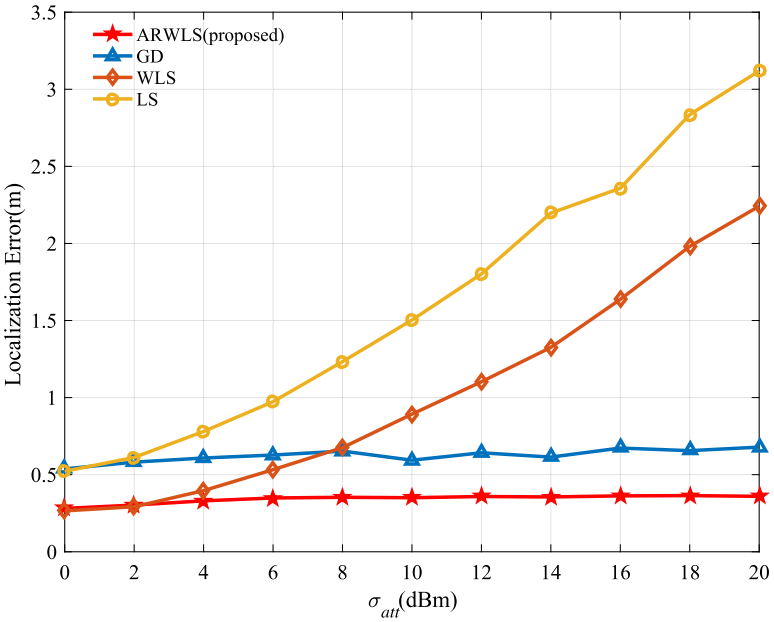
Fig. 4. Localization error with varied std deviation of attacks, σ_{att}, at different levels of noises, σ.

and improves the robustness in the location calculation stage. When the attack strength (σ_{att}) increases, the proposed scheme will introduce large errors to the localization results once a malicious anchor node participates in the location calculation stage. At the same time, it is more likely that the power difference samples of the malicious anchor node exceed the reference error interval. Therefore, the probability of the malicious anchor nodes being eliminated also increases. Accordingly, the fluctuation range of the localization error becomes relatively small.

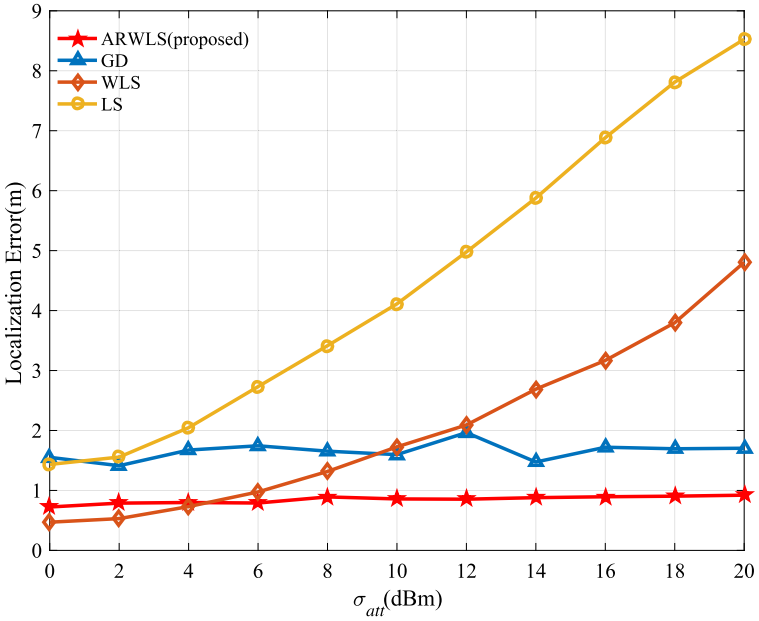
In contrast, the information provided by anchor nodes is taken into account for the WLS algorithm. The participation of malicious anchor nodes in localization will inevitably affect the performance of the localization mechanism. For the GD algorithm, it directly eliminates 50% of anchor nodes, leading to a part of benign anchor nodes unable to participate in the final localization. As is shown in Fig. 4(b), when $\sigma > \sigma_{att}$, the localization error of the proposed algorithm is slightly larger than the error of the WLS algorithm. However, when σ_{att} continues increasing, the localization error of the proposed algorithm keeps lower than 2 m till $\sigma_{att} = 20$ dBm, while the error for other three algorithms grows much faster.

In the above experiments, the location of the target node is deployed randomly each time. For the fixed location of the target node, Fig. 5 shows the localization performance of the algorithms. Figure 5(a) shows the performance of the localization error of the algorithms with the standard deviation of the attack σ_{att} when the target node is fixed at the center of the deployment area, (50, 50) and the anchor nodes are randomly distributed, while Fig. 5(b) shows that when the target node is fixed at the edge of the area, (10, 90). Compared with the curves in Fig. 5, the localization error of the proposed ARWLS algorithm is always kept below others, specifically, below 0.5 m and 1 m when the location of the target node is fixed at the center or near the boundary of the deployment area, respectively. Further, compared with other three algorithms, the ARWLS algorithm is less affected as the attack strength increases, demonstrating a strong robustness.

Coordinated Attacks. The case of coordinated attacks is rather complicated. To explain the influence of coordinated attacks, we provide experiments to explain. In our experiments, the percentage of the malicious anchor nodes is 30%. Figure 6 compares the localization performance of the algorithms with a varying distance between the real location of the target node $\mathbf{T}$ and the shifted location, $\mathbf{T}_{mal}$, tampered by the attacker under coordinated attacks. It is observed that the performance of the WLS is similar to that of the LS. Such a result means that the WLS does not greatly improve the security of the network under coordinated attacks. When $d_a < 20$ m, the localization error of the proposed algorithm is slightly higher than the GD algorithm. This is because the accuracy of anchor node screening stage of the proposed algorithm is relatively low when the value of d_a is small. However, when d_a continues increasing, the proposed algorithm outperforms others, which shows the excellence of the proposed ARWLS algorithm most of the time.



(a) Localization error with the target node fixed at (50, 50), the center of the deployment area



(b) Localization error with the target node fixed at (10, 90), the edge of the area

Fig. 5. Localization error with varied std deviation of attacks, σ_{att}, under noncoordinated attacks.

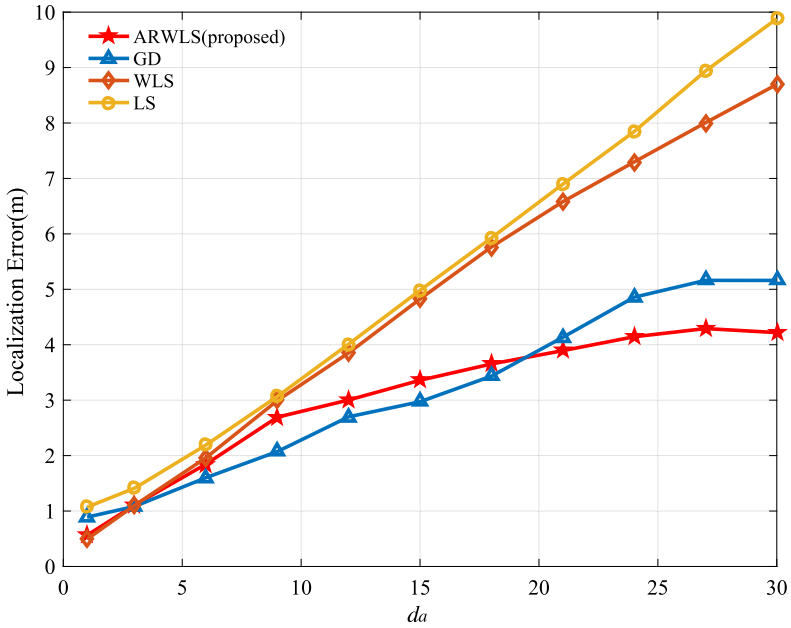


Fig. 6. Localization error with the distance between the real location of the target node and the shifted location d_a under coordinated attacks.

5 Conclusion

Malicious anchor nodes may tamper transmitting power and launch non-coordinated or coordinated attack to the network, which badly affects the estimate accuracy in the RSSI-based localization mechanism. In order to solve the problem of node localization in such a malicious environment, a localization algorithm named ARWLS is proposed in this paper. The algorithm can be implemented without requiring prior information or additional hardware support. The proposed ARWLS algorithm is performed in two stages: the anchor nodes screening stage and the location calculation stage. In the algorithm, the quasi-Newton method is used to determine the reference anchor nodes and the initial location estimate. The approximately calculated power difference is obtained and used as the sample of the SPRT for further screening. Finally, the weighted least squares method is used to improve the robustness of location calculation. The proposed algorithm is compared with the existing algorithms in the performance of localization accuracy. The simulation results show that the proposed algorithm is superior to others, especially in the case of the non-coordinated attacks.

References

1. Khan, I., Belqasmi, F., Glitho, R., Crespi, N., Morrow, M., Polakos, P.: Wireless sensor network virtualization: a survey. IEEE Commun. Surv. Tutor. **18**(1), 553–576 (2016)
2. Liu, X., Li, W., Han, F., Xie, Y.: An optimization scheme of enhanced adaptive dynamic energy consumption based on joint network-channel coding in WSNs. IEEE Sens. J. **17**(18), 6119–6128 (2017)
3. Lu, W., Gong, Y., Liu, X., Wu, J., Peng, H.: Collaborative energy and information transfer in green wireless sensor networks for smart cities. IEEE Trans. Ind. Informat. **14**(4), 1585–1593 (2018)
4. Zhang, Y., Sun, L., Song, H., Cao, X.: Ubiquitous WSN for healthcare: recent advances and future prospects. IEEE Internet Things J. **1**(4), 311–318 (2014)
5. Zhou, B., Chen, Q.: On the particle-assisted stochastic search mechanism in wireless cooperative localization. IEEE Trans. Wirel. Commun. **15**(7), 4765–4777 (2016)
6. Liu, D., Xu, Y., Huang, X.: Identification of location spoofing in wireless sensor networks in non-line-of-sight conditions. IEEE Trans. Ind. Informat. **14**(6), 2375–2384 (2018)
7. Liu, X., Yin, J., Zhang, S., Ding, B., Guo, S., Wang, K.: Range-based localization for sparse 3-D sensor networks. IEEE Internet Things J. **6**(1), 753–764 (2019)
8. Sun, Y., Zhang, F., Wan, Q.: Wireless sensor network-based localization method using TDOA measurements in MPR. IEEE Sens. J. **19**(10), 3741–3750 (2019)
9. Xiong, H., Peng, M., Gong, S., Du, Z.: A Novel hybrid RSS and TOA positioning algorithm for multi-objective cooperative wireless sensor networks. IEEE Sens. J. **18**(22), 9343–9351 (2018)
10. Wu, Y.I., Wang, H., Zheng, X.: WSN localization using RSS in three-dimensional space-a geometric method with closedform solution. IEEE Sens. J. **16**(11), 4397–4404 (2016)
11. Zhao, Y., Liu, X., Han, F., Han, G.: Recovery of hop count matrices for the sensing nodes in Internet of Things. IEEE Internet Things J. **7**(6), 5128–5139 (2020)
12. Ahmadi, Y., Neda, N., Ghazizadeh, R.: Range free localization in wireless sensor networks for homogeneous and non-homogeneous environment. IEEE Sens. J. **16**(22), 8018–8026 (2016)
13. Fan, J., Hu, Y., Luan, T.H., Dong, M.: DisLoc: a convex partitioning based approach for distributed 3-D localization in wireless sensor networks. IEEE Sens. J. **17**(24), 8412–8423 (2017)
14. Liu, X., Su, S., Han, F., Liu, Y., Pan, Z.: A range-based secure localization algorithm for wireless sensor networks. IEEE Sens. J. **19**(2), 785–796 (2019)
15. Garg, R., Varna, A.L., Wu, M.: An Efficient gradient descent approach to secure localization in resource constrained wireless sensor networks. IEEE Trans. Inf. Forens. Secur. **7**(2), 717–730 (2012)
16. Liu, D., Ning, P., Wenliang, D.: Detecting malicious beacon nodes for secure location discovery in wireless sensor networks. In: 25th IEEE International Conference on Distributed Computing Systems (ICDCS 2005), pp. 609–619, Columbus (2005)
17. Brands, S., Chaum, D.: Distance-bounding protocols. In: Helleseth, T. (ed.) EUROCRYPT 1993. LNCS, vol. 765, pp. 344–359. Springer, Heidelberg (1994). https://doi.org/10.1007/3-540-48285-7_30
18. Li, Z., Trappe, W., Zhang, Y., Nath, B.: Robust statistical methods for securing wireless localization in sensor networks. In: Fourth International Symposium Information Processing in Sensor Networks, pp. 91–98, April 2005

19. Mukhopadhyay, B., Srirangarajan, S., Kar, S.: Robust range-based secure localization in wireless sensor networks. In: IEEE Global Communications Conference (GLOBECOM). Abu Dhabi, United Arab Emirates 2008, pp. 1–6 (2018)
20. Liu, D., Ning, P., Du, W.K.: Attack-resistant location estimation in sensor networks. In: IPSN: Fourth International Symposium on Information Processing in Sensor Networks, 2005, Boise, ID, USA 2005, pp. 99–106 (2005)
21. Capkun, S., Rasmussen, K., Cagalj, M., Srivastava, M.: Secure location verification with hidden and mobile base stations. IEEE Trans. Mob. Comput. **7**(4), 470–483 (2008)
22. Capkun, S., Hubaux, J.: Secure positioning in wireless networks. IEEE J. Sel. Areas Commun. **24**(2), 221–232 (2006)
23. Du, W., Fang, L., Ningi, P.: LAD: localization anomaly detection for wireless sensor networks. In: 19th IEEE International Parallel and Distributed Processing Symposium, pp. 1–15, CO, Denver (2005)
24. Jin, L., Zhang, Y.: Discrete-time zhang neural network for online time-varying nonlinear optimization with application to manipulator motion generation. IEEE Trans. Neural Netw. Learn. Syst. **26**(7), 1525–1531 (2015)
25. Rappaport, T.S.: Wireless Communications: Principles and Practice. Prentice-Hall, Upper Saddle River (2002)
26. Koh, J.Y., Nevat, I., Leong, D., Wong, W.: Geo-spatial location spoofing detection for Internet of Things. IEEE Internet Things J. **3**(6), 971–978 (2016)
27. Zhang, P., Nagarajan, S.G., Nevat, I.: Secure location of things (SLOT): mitigating localization spoofing attacks in the Internet of Things. IEEE Internet Things J. **4**(6), 2199–2206 (2017)
28. Ho, J., Wright, M., Das, S.K.: ZoneTrust: fast zone-based node compromise detection and revocation in wireless sensor networks using sequential hypothesis testing. IEEE Trans. Dependable Secure Comput. **9**(4), 494–511 (2012)