

Research Article

Heuristic Localization Algorithm with a Novel Error Control Mechanism for Wireless Sensor Networks with Few Anchor Nodes

Yujia Sun , Xiaoming Wang , Jiyan Yu, and Yu Wang

Ministerial Key Laboratory of ZNDY, Nanjing University of Science and Technology, Nanjing 210094, China

Correspondence should be addressed to Xiaoming Wang; 202xm@163.com

Received 11 November 2017; Revised 23 May 2018; Accepted 7 June 2018; Published 11 July 2018

Academic Editor: Bruno Andò

Copyright © 2018 Yujia Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A novel iterative localization algorithm with high accuracy and low anchor node dependency for large-scale wireless sensor networks is proposed in this paper. At each iteration, blind nodes are located using a weighted linear least squares-based algorithm. To prevent errors in the blind nodes from propagating and accumulating throughout the network, an anchor geometric feature-based error control mechanism is used to select the nodes that participate in the localization and to estimate the localization confidence. The simulation results show that the algorithm can be used when only a few anchor nodes are involved. This algorithm is more advanced than traditional methods, which often require a large number of well-placed anchor nodes to operate appropriately. By optimizing the decision parameter ν of the algorithm, the average localization error of the algorithm is approximately 0.43 meters. When the ratio of anchor nodes (the ratio of the number of anchor nodes to the number of sensor nodes in the network) is 1.25% (i.e., 5 anchor nodes for 400 sensor nodes), the received signal strength indicator (RSSI) variance is 8 dBm, and the radio range is 50 meters. A comparison of the proposed algorithm with global localization methods, including multidimensional scaling (MDS), semidefinite programming (SDP), and shortest-path access (SPA), shows that the proposed algorithm achieves higher location accuracy and stability when the number of anchor nodes is varied. The efficiency of the proposed localization algorithm is evaluated in a real sensor network, and the accuracy is high and robust to radio channel variance.

1. Introduction

Wireless sensor networks (WSNs) are a basic component of the Internet of Things. Based on the development of mobile computing and embedded technologies, WSNs have been widely implemented in daily applications, such as health care monitoring, natural disaster prevention, and surveillance [1–3]. Sensor network services, such as mobile sinks, geographic routing, and location-based multicasting, rely on the physical location information of sensor nodes or the phenomena of interest. One method of accessing the locations of nodes is to use the Global Positioning System (GPS). However, installing a GPS chip on every sensor node is expensive. Moreover, the GPS is not always available because of environmental or hardware limitations, such as shadowing, energy, volume, or cost issues. Therefore, efficient and inexpensive localization technologies must be developed [4, 5].

Various WSN localization approaches have been presented in the literature. Depending on the usage of the range information between nodes, node localization approaches can be classified into range-based localization and range-free localization. In a range-based localization method, the blind nodes estimate the distance to each anchor node and then estimate the location based on the distance to and position of the anchor nodes. Anchor nodes are a special type of sensor node with known locations, with the help of GPS, or are predeployed. Blind nodes are sensor nodes with unknown positions. To estimate distance, blind nodes use radio channel information, such as the received signal strength indicator (RSSI) [6, 7], time of arrival (ToA) [8, 9], time difference of arrival (TDoA) [10, 11], angle of arrival (AoA) [12], or some combination of these methods. The RSSI-based distance estimation is attractive because of its low cost, long range, and simplicity. After determining the distance between blind

nodes and anchor nodes, multilateration technology is used for position estimation [13, 14]. Trilateration is a particular form of multilateration that utilizes three anchor nodes to calculate the location of a blind node in two dimensions. For a given WSN, to estimate the position of all blind nodes, at least three anchor nodes are required in the communication range of the blind node. For a large-scale outdoor randomly deployed WSN, a large number of anchor nodes should be deployed to guarantee that every blind node is located. However, additional anchor nodes correspond to additional costs associated with hardware, deployment, maintenance, and so on.

Iterative node localization is a novel localization method that can decrease the dependence on anchor nodes by importing a blind node leveling-up scheme [15, 16]. For iterative node localization, the located blind node is upgraded to new anchor nodes (pseudoanchor) in which the location information of the initial anchor nodes is progressively propagated to other nodes. Theoretically, all blind nodes in a network can be located with three anchor nodes under an iterative localization method. In practical localization systems, measurement noise during node ranging is inevitable, and the location of a pseudoanchor is uncertain. As other blind nodes locate themselves and refer to the pseudoanchors with uncertain locations, the localization error increases as more pseudoanchors are introduced into the localization. This process is called the error propagation problem in iterative node localization [17]. Various methods have been proposed to mitigate the propagation error. Liu et al. [18] proposed an error control mechanism that tracks the estimation error and quantifies each location estimation with a given level of uncertainty. Based on the location certainty information, the algorithm introduces a neighbor selection procedure and an update criterion. The simulation results show that the localization accuracy and robustness are significantly improved with this approach. Yang and Liu [19] proposed the concept of the quality of trilateration (QoT), which considers both the geometric relationship and ranging errors. Based on the QoT, a confidence-based iterative localization (CIL) scheme can be designed. At each stage, the CIL selectively utilizes trilaterations and reduces the likelihood of using low-confidence references, which effectively halts error propagation. Wu et al. [20] considered the error control problem for non-Gaussian noise measurements. The proposed error control algorithm estimates the location error based on nonlinear least squares residuals. Additionally, a robust formulation of error control is proposed that can reduce the accumulated error. The simulations show that more anchors, less noise, and fewer iterations reduce the localization error. Hu et al. [21] derived an upper bound of error propagation for iterative localization that works well when the precise environmental noise distribution is unavailable. The minimum upper bound is adopted to evaluate the localization result. With this method, the algorithm constructs the proper linear least squares localization with high probability. The main concept underlying these methods is the estimation of the ranging error. The algorithms exhibit good localization accuracy when the noise is precisely estimated.

In this paper, we analyze the localization problem to reduce anchor node dependency in a large-scale WSN. A novel iterative localization algorithm with an anchor geometric feature-based error control mechanism is proposed. Compared with previous research that precisely modeled the localization error during each iteration, we formulate the error control problem as a classification problem. Based on node localization data under varying ratios of anchor nodes, RSSI variances, and radio communication ranges in a randomly deployed node environment, we characterize the localization condition by considering the number of anchor nodes, the anchor positions, the distance between the anchors and blind nodes, and the spatial distribution of anchors. Additionally, a one-class support vector decision-maker is trained based on the selected features. The simulation results show that the algorithm can be used when only a few anchor nodes are involved, which is more advanced than traditional methods that often require a large number of well-placed anchor nodes to work well. We evaluate the performance of our algorithm with different numbers of anchor nodes and RSSI variances. The results show that the algorithm efficiently solves the error propagation problem. The average localization error of the algorithm is approximately 0.43 meters when the percentage of anchor nodes is 1.25%, the RSSI variance is 8 dBm, and the radio range is 50 meters. The location accuracy is better than those of certain global localization methods, including MDS, SDP, and SPA. The efficiency of the proposed localization algorithm is evaluated in a real sensor network, and the accuracy is high and robust to radio channel variance.

The remainder of this paper is organized as follows. In Section 2, we present the RSSI channel model and the weighted linear least squares location estimator. In Section 3, we present the anchor geometric features and the proposed error control mechanism. In Section 4, we evaluate the proposed localization algorithm with simulation experiments. In Section 5, we conclude the paper.

2. Localization Method

In this paper, we classify wireless sensor nodes as blind nodes, anchor nodes, pseudoanchor nodes, and reference nodes based on their localization duty. These names will be discussed in the following sections, and their respective duties are explained at the beginning of the paper. Anchor nodes are nodes that can access their absolute location at the beginning of the deployment. Blind nodes are nodes that do not know their location and attempt to estimate it based on the locations of the anchor nodes. A pseudoanchor node is a blind node that has an estimated location and is upgraded to an anchor node. The location of the pseudoanchor node always has an error, and the algorithm should be designed to stop the error from propagating throughout the network. The reference node is a node used by the blind node to locate itself. Both the anchor nodes and the pseudoanchor nodes can be used as reference nodes.

2.1. RSSI Signal Model. Many channel models have been proposed in the literature for indoor or outdoor localization

applications [22]. In this paper, the log-normal shadowing model is used for RSSI-based localization because of its simplicity. Let P_{RX} denote the received power and d denote the distance between the transmitter and the receiver. Then, the log-normal shadowing model is expressed as follows:

$$P_{RX}(\text{dBm}) = A - 10\eta \log \frac{d}{d_0} + N, \quad (1)$$

where A is the path loss at distance d_0 , η is the path loss exponent, and $N \sim N(0, \sigma^2)$ is a zero mean random variable with variance σ . In general, d_0 is equal to 1 meter, and η is in the range of 2 to 4. The parameters A and η should be measured through real experiments. In this paper, we conduct the experiment based on the TI CC2650 sensor node in an outdoor environment. The results are shown in Figure 1. This channel model is used in the following section.

2.2. Weighted Linear Least Squares Estimator. The distance \tilde{d}_i between a blind node and each reference i ($i = 1, 2, \dots, m$) is estimated based on the radio channel model. Let (x, y) represent the position of the blind node and (x_i, y_i) represent the position of the reference. The distance estimation error is expressed as follows:

$$\varepsilon = \sum_{i=1}^m \left(\sqrt{(x_i - x)^2 + (y_i - y)^2} - d_i \right)^2. \quad (2)$$

The localization problem finds the optimal (x, y) that minimizes ε . Formula (2) is a nonlinear function, and the optimal position is iteratively calculated using a searching method, such as a gradient search, as follows:

$$\begin{bmatrix} x \\ y \end{bmatrix}_{k+1} = \begin{bmatrix} x \\ y \end{bmatrix}_k - \alpha \begin{bmatrix} \frac{\partial \varepsilon}{\partial x} \\ \frac{\partial \varepsilon}{\partial y} \end{bmatrix}_{x=x_k, y=y_k}. \quad (3)$$

Searching methods require an initial estimate of the position, which is difficult in real environmental applications. Therefore, we prefer to translate the problem into a linear problem that can be solved with a least squares estimator. Based on the positions of the blind node and the references, the distance between them can be expressed as follows:

$$d_i^2 = (x_i - x)^2 + (y_i - y)^2, \quad (4)$$

where d_i is the real distance, which is different from the noisy estimation distance \tilde{d}_i . By setting reference node 1 as an origin, we obtain $d_1^2 = (x_1 - x)^2 + (y_1 - y)^2$. Then, for the other references, we obtain the following equation:

$$d_i^2 - d_1^2 = (x_i - x)^2 + (y_i - y)^2 - (x_1 - x)^2 - (y_1 - y)^2, \quad (5)$$

where $i > 1$. Formula (5) can be expressed in matrix form as follows:

$$\mathbf{G} - 2\mathbf{W}\mathbf{P} = 0, \quad (6)$$

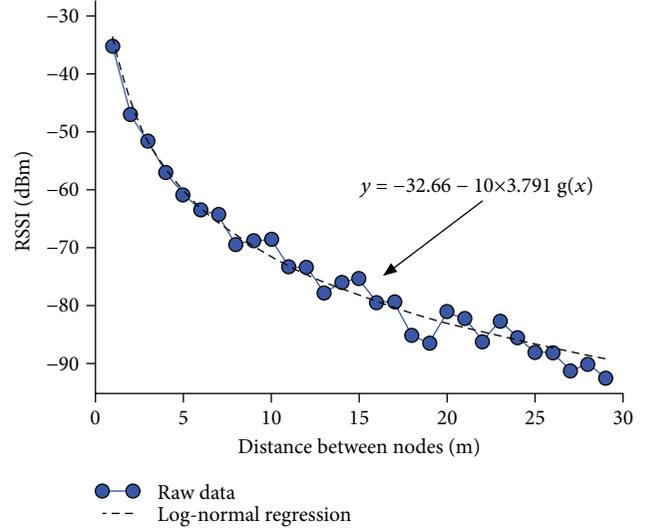


FIGURE 1: RSSI model in the outdoor environment.

where

$$\mathbf{G} = \begin{bmatrix} d_2^2 - d_1^2 - x_2^2 - y_2^2 + x_1^2 + y_1^2 \\ d_3^2 - d_1^2 - x_3^2 - y_3^2 + x_1^2 + y_1^2 \\ \vdots \\ d_m^2 - d_1^2 - x_m^2 - y_m^2 + x_1^2 + y_1^2 \end{bmatrix}, \quad (7)$$

$$\mathbf{W} = \begin{bmatrix} x_1 - x_2 & y_1 - y_2 \\ x_1 - x_3 & y_1 - y_3 \\ \vdots & \vdots \\ x_1 - x_m & y_1 - y_m \end{bmatrix}.$$

In a real RSSI-based localization system, the real distance between nodes is never obtained. Instead, a noisy distance estimation \tilde{d}_i is obtained. Then, matrix \mathbf{G} can be formulated as follows:

$$\tilde{\mathbf{G}} = \begin{bmatrix} \tilde{d}_2^2 - \tilde{d}_1^2 - x_2^2 - y_2^2 + x_1^2 + y_1^2 \\ \tilde{d}_3^2 - \tilde{d}_1^2 - x_3^2 - y_3^2 + x_1^2 + y_1^2 \\ \vdots \\ \tilde{d}_m^2 - \tilde{d}_1^2 - x_m^2 - y_m^2 + x_1^2 + y_1^2 \end{bmatrix}. \quad (8)$$

Now, the position of the blind node is calculated as the least squares solution of (6), which is expressed as follows:

$$\mathbf{P} = \frac{1}{2} (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \tilde{\mathbf{G}}. \quad (9)$$

The RSSI signal is sensitive to environmental interference. The linear least squares localization solution exhibits poor accuracy when the RSSI variance is increasing. To improve the localization accuracy, a weighted linear least squares estimator is introduced, in which the solution is weighted based on the variance of the distance estimation.

The weighted linear least squares solution is expressed as follows:

$$P = \frac{1}{2} (\mathbf{W}^T \mathbf{S}^{-1} \mathbf{W})^{-1} \mathbf{W}^T \mathbf{S}^{-1} \tilde{\mathbf{G}}, \quad (10)$$

where matrix \mathbf{S} is the covariance matrix of $\tilde{\mathbf{G}}$. Assuming that the positions of the references are constant and that the distance measurements are independent, matrix \mathbf{S} is expressed as follows:

$$\mathbf{S} = \begin{bmatrix} v(\tilde{d}_1^2) + v(\tilde{d}_2^2) & v(\tilde{d}_1^2) & \cdots & v(\tilde{d}_1^2) \\ v(\tilde{d}_1^2) & v(\tilde{d}_1^2) + v(\tilde{d}_3^2) & \cdots & v(\tilde{d}_1^2) \\ \vdots & \vdots & \ddots & \cdots \\ v(\tilde{d}_1^2) & v(\tilde{d}_1^2) & \cdots & v(\tilde{d}_1^2) + v(\tilde{d}_m^2) \end{bmatrix}, \quad (11)$$

where $V()$ is the variance of each distance measurement.

The weighted linear least squares localization displays increased accuracy and robustness and less complexity than traditional methods.

3. Error Control Method

3.1. Geometric Features of Localization. Multilateration is infeasible when three references are collinear, and it is feasible when three references are not collinear [21]. In a large-scale randomly deployed WSN, the references are rarely collinear. However, this does not mean that the blind node has a good localization result for any type of anchor node position. Therefore, we conduct a simulation to investigate the relationship between the reference node geometry and localization accuracy. In the simulation, the network includes 400 wireless sensor nodes that are randomly deployed in a square area with an edge equal to 200 m. The blind node is located based on (10), and the RSSI measurement suffers from Gaussian noise with a mean of zero and a variance of 2 dBm. Additional simulation details are presented in Table 1.

Based on the localization results of the simulation, six typical localization situations for a single blind node with different numbers of anchor nodes are plotted in Figure 2. The red triangle represents the position of the blind node, and the blue dots and yellow stars represent the positions of the anchor nodes and the estimated positions of the blind node, respectively. The red circle is the communication range of the blind node. Figures 2(a)–2(c) are the situations where the blind node is associated with a bad localization result, with numbers of anchor nodes equal to 6, 10, and 16, respectively. The distribution of the estimated positions becomes concentrated when there are more anchor nodes, which means that the localization accuracy improves when the number of anchor nodes increases. Figures 2(d)–2(f) are situations in which the blind node is associated with a good localization result, with the same number of anchor nodes. Compared with Figure 2(a), the localization result is much more accurate in Figure 2(d). As the number of anchor nodes

TABLE 1: Description of simulation parameters.

Parameter	Description	Default
M	Number of blind nodes	380
N	Number of anchor nodes	20
L	Number of initial pseudoanchors	5
R	Radio communication range	30 m
σ	RSSI variance	1 dBm
A	Path loss at one meter	32.6 dBm
η	Path loss exponent	3.79
ϕ	OC-SVM kernel type	Linear
ν	OC-SVM parameter nu	0.1
e	OC-SVM parameter epsilon	0.001

and the network configuration are the same for the situations in Figures 2(a) and 2(d), the geometry of the anchor nodes in Figure 2(d) can help improve the localization accuracy. Furthermore, a comparison of Figures 2(c) and 2(d) reveals that this improvement is significant. Figure 2(d) is more accurate than Figure 2(c) with a smaller number of anchor nodes (6 in Figure 2(d) and 16 in Figure 2(c)). In this paper, offset and collinear are used to describe the geometric pattern of the anchor nodes, and the quantified features are discussed in Section 3.2. The pattern offset means that the centroid of the anchor nodes is far from the blind node, as Figures 2(b) and 2(c) show. In this situation, the estimated position tends to be in the direction of the vector centroid based on the real blind node position. The collinear pattern is not a real collinear condition, but the anchor nodes are almost located in a single line, as Figure 2(a) shows. In this pattern, the centroid of the anchor node may be close to the real position of the blind node, but the estimated position has a large variance in the direction from the centroid to the real blind node position.

As Figure 2 shows, the number of anchor nodes affects the localization accuracy of the blind node, and the result is more accurate when the number of anchor nodes is high. We collect the number of references of each blind node and the associated localization error during the simulation. The data are plotted in Figure 3. The number of references has a large range from 3 to 25. Additionally, for most blind nodes, the number of references ranges from 10 to 15. The localization error decreases as the number of references increases. Moreover, the localization results exhibit considerable uncertainty when the number of references is less than 5, as the anchor nodes tend to display a bad geometric pattern. In this way, the number of references can be used as a feature for predicting localization errors that are larger than a specified threshold.

Our observations show that the number of references and the reference geometric pattern significantly affect the localization accuracy. When the references are not evenly distributed around the blind node or when the number of references is less than the threshold, the localization results will exhibit considerable uncertainty. Therefore, we believe

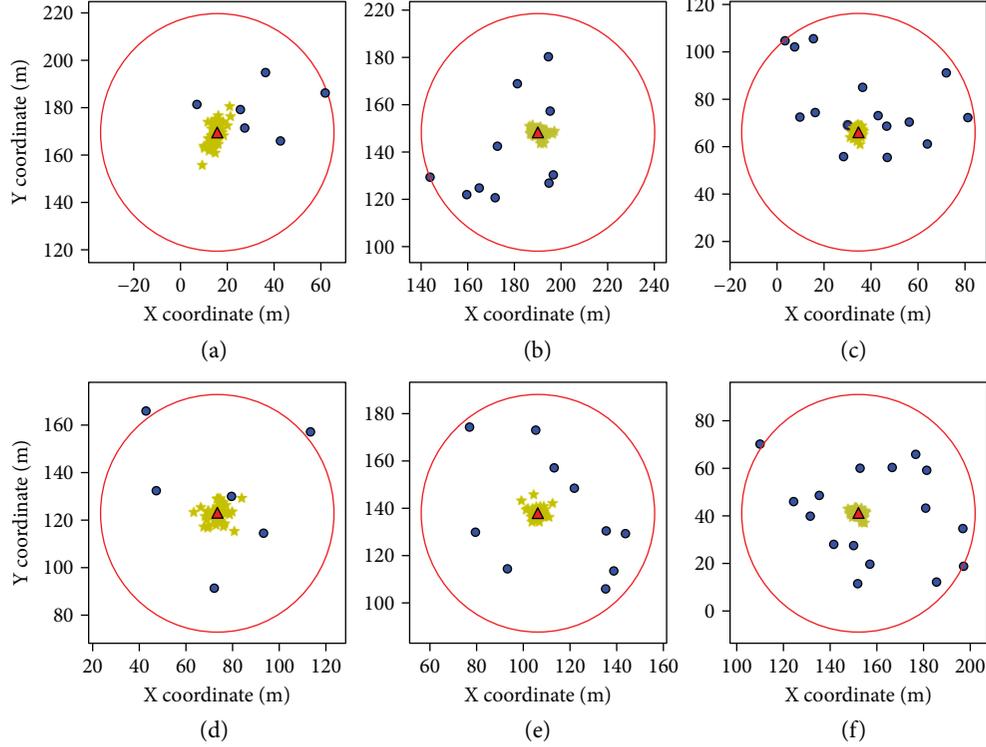


FIGURE 2: Estimation layouts for different anchor (neighbor) geometries. (a) Uncertain with 6 references; (b) uncertain with 10 references; (c) uncertain with 16 references; (d) stable with 6 references; (e) stable with 10 references; and (f) stable with 16 references.

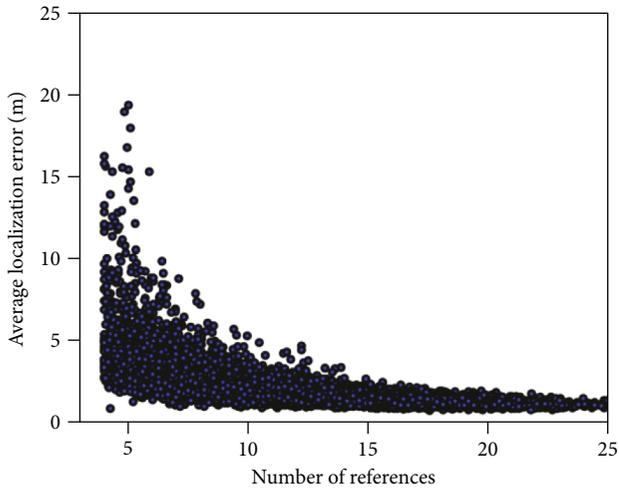


FIGURE 3: Localization error for different numbers of references.

that the localization error can be predicted by characterizing the reference geometry.

3.2. Feature Selection. The uncertain locations of the references, the distance estimation errors, and the reference geometry are the factors that affect the localization accuracy. These features can be used to predict the localization accuracy and inform the decision-making in the blind upgrading process.

The covariance matrix is used to describe the data distribution in a higher dimension, which is an extension of the variance. In our implementation, we construct the covariance matrix based on the coordinates of the references in a 2-dimensional plane. Given a blind node, let $X = \{x_1, x_2, \dots, x_m\}$ denote the x coordinates of the references and $Y = \{y_1, y_2, \dots, y_m\}$ denote the y coordinates of the references, where m represents the number of references. The covariance matrix of X and Y is given as follows:

$$C = \begin{bmatrix} \text{cov}(X, X) & \text{cov}(X, Y) \\ \text{cov}(Y, X) & \text{cov}(Y, Y) \end{bmatrix}, \quad (12)$$

where $\text{cov}(X, Y) = E[(X - (EX))(Y - (EY))]$, and $E(X)$ is the expectation of X . Our geometric features are based on the eigenvalue of C . We denote the larger eigenvalue as λ_1 and the smaller one as λ_2 . Then, we define a feature ALC to describe the linearity of a set of references because this linearity causes unstable localization. ALC can be expressed as follows:

$$\text{ALC} = \frac{\lambda_1 - \lambda_2}{\lambda_1}. \quad (13)$$

When the spatial location of a set of references is located on a line, ALC is equal to 1. However, when the set of references is located on a circle, ALC is equal to 0.

To describe the uniformity of the references, we define two features ALR and ALS. ALR is expressed as follows:

$$\text{ALR} = \left| \frac{R^2 - \lambda_1 \lambda_2}{R^2(d \max - d \min)} \right|. \quad (14)$$

ALS is expressed as follows:

$$\text{ALS} = \frac{\sum_{i=1}^m \left| \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2} - d_i \right|}{m}, \quad (15)$$

$$(x_c, y_c) = \left(\frac{\sum_{i=1}^m x_i}{m}, \frac{\sum_{i=1}^m y_i}{m} \right), \quad (16)$$

where (x_c, y_c) represents the centroid of the references, R is the communication range of the blind node, and d_i is the distance from the blind node to each reference. When the references are uniformly located around the blind node, the values of ALR and ALS tend to be 0; however, the values here are 1 and R , respectively.

In addition to the previously proposed features, several basic data features used in classification applications are shown in Table 2. These basic features work on three signal channels, which are the X coordinates of the references, the Y coordinates of the references, and the distance measurement from each reference. Based on these features, a feature space is constructed, and the importance of the features is analyzed in the following section.

In total, 25 features are extracted, but some are irrelevant or redundant features. Therefore, feature selection and feature ranking are applied to identify the important features in the feature space. Feature selection methods can be classified as the wrapper method, the embedded method, and the filter method. In this paper, we apply a two-stage feature selection procedure. First, the maximal information coefficient (MIC) is used to score each feature. Then, the 10 best features are selected to form a feature subset, and a recursive feature elimination (RFE) method is used to select the 4 best features from the subset. The MIC is a measure of the strength of the linear relationship between two variables X and Y . The basic concept of the MIC is to use the relevant mutual information (MI). To obtain the MIC score, the values of two variables are separated into different numbers of bins to construct rectangular grids with different resolutions. Thus, the distribution of the cells in each grid is the fraction of points that fall into that cell. Then, the MI statistics for each grid are calculated, and the maximum is chosen as the MIC score. The MIC of two variables X and Y is defined as follows:

$$\text{MIC}(X, Y) = \max_{x_n, y_n < N^{0.6}} \left[\frac{\text{MI}_{x_n, y_n}(x, y)}{\log_2 \min \{x_n, y_n\}} \right], \quad (17)$$

where N is the sample size; x_n and y_n denote the number of bins for the x - and y -axes, respectively; and MI_{x_n, y_n} are the MI statistics between the variables for an $x_n - by - y_n$ rectangular grid. We calculate the MIC between the feature and the localization error based on (17). The result is used to score the corresponding feature, as shown in Figure 4.

The 10 best features with their corresponding ranks are shown in Table 3. The features proposed in this paper yield

TABLE 2: Signal features from the literature.

Feature	Description
Len()	Number of references
Mean()	Mean value
Std()	Standard deviation
Max()	Largest value in a series
Min()	Smallest value in a series
Sma()	Signal magnitude area
Energy()	Sum of squares divided by the number of values
Entropy()	Signal entropy

high MIC scores and are ranked 1, 2, and 4. In addition, the feature pLen is ranked 3, which means that the number of references is highly related to the localization error. For all possible combinations of up to four of these features, we test the classification accuracy based on a support vector machine (SVM). Table 4 details the accuracy of the three best and the three worst feature combinations. The feature combination with the most accurate results (ALC, ALR, ALS, and pLen) is used for further analysis.

3.3. Error Control Algorithm. As previously noted, in a large-scale WSN localization problem, the percentage of anchor nodes is expected to be small, which results in a situation in which a small percentage of blind nodes can directly communicate with the anchor nodes. However, none of the blind nodes in the network may be able to find at least three anchor nodes to derive a location. To overcome this problem, a two-phase localization algorithm that includes the initialization phase and the error control phase is proposed.

Considering the small number of anchor nodes, the initialization algorithm will find a series of pseudoanchor nodes and estimate their positions based on a distance vector-hop (DV-hop) [23] distance estimation. Initially, the anchor nodes in the network broadcast a packet that contains the unique ID and position of the node and hop information (equal to 0). All the neighbors surrounding the anchor node receive the packets. The RSSI, position of the anchor node, and hop information of 1 are recorded by the one-hop neighbors, and they then broadcast another packet containing the anchor position and hop information of 1. The nodes that receive the second packet and did not receive the first packet will record the anchor position and hop information of 2, and they then begin a new broadcast. The anchor nodes that receive a packet with hop information of 1 record the unique ID and the RSSI value of the packet to form a one-hop neighbor list. This process continues until each anchor position and the associated hop information have been spread to every node in the network. The nodes that receive hop information multiple times search for the smallest number of hops for each anchor node. This process improves the accuracy of a DV-hop distance estimation and reduces network traffic.

By determining the hop information associated with each anchor node, a blind node estimates the distances by multiplying the hop information by a shared average hop distance metric. The DV-hop distance estimation suffers from large

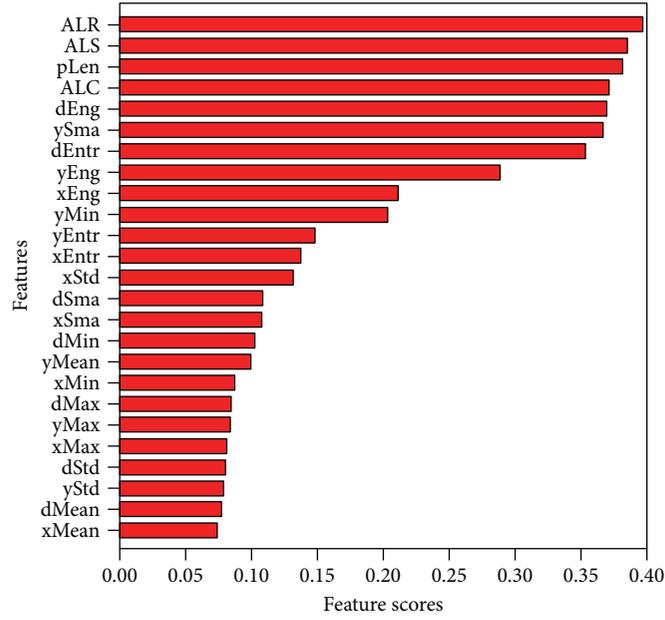


FIGURE 4: MIC scores for 25 different features.

TABLE 3: Rank of the 10 best features.

Rank	1	2	3	4	5
Name	ALR	ALS	pLen	ALC	dEng
Rank	6	7	8	9	10
Name	ySma	dEntr	yEng	xEng	yMin

TABLE 4: Classification accuracy for different feature combinations.

Accuracy	1	2	3	4	5	6	7	8	9	10
0.843	×	×	×	×						
0.839		×		×		×	×			
0.835	×	×	×			×				
0.636							×	×	×	×
0.645					×	×			×	×
0.649						×	×	×		×

error compared with those of channel model-based methods because the localization error of the DV-hop estimation process is large. Therefore, multiple blind nodes are selected as pseudoanchor nodes in the initialization phase. The blind nodes in the one-hop range of each anchor node are the best choices of pseudoanchor nodes because one reference distance can be estimated based on the channel model. Given the one-hop neighbors of an anchor, the blind nodes are sorted by their distance to the anchor node, and the nearest n_{init} blind nodes are the pseudoanchor nodes. Then, the positions of these pseudoanchor nodes are estimated with (10). The initialization phase is summarized in Algorithms 1 and 2.

In an iterative localization scenario, the position of a pseudoanchor node is uncertain. This positional error may further affect the localization of the other nodes and increase

Input: Number of initial Pseudo-anchor n_{init}
Define: HopInfo Packet $HIP = \{ID, position, hopInfo\}$
 NeighborInfo Packet $NIP = \{ID_1, \dots, ID_{n_{init}}\}$
 Neighbor RSSI List $NRL = [ID_i, RSSI_i]$
 Neighbor Distance List $NDL = [ID_i, d_i]$

Procedures:
 1: send an HIP
 2: **while** not time up:
 3: **if** received an HIP :
 4: parse the HIP to obtain $ID, position, hopInfo$
 5: obtain the $RSSI$
 6: **if** $hopInfo$ is equal to 1:
 7: append $[ID, RSSI]$ to NRL
 8: reset the timer
 9: **else:**
 10: pass
 11: **end while**
 12: **for** i from 1 to $sizeof(NRL)$:
 13: estimate d_i based on $RSSI_i$ by Eq. (1)
 14: append $[ID_i, d_i]$ to NDL
 15: **endfor**
 16: sort NDL by d_i in ascending
 17: **if** $sizeof(NDL) > n_{init}$:
 18: set NIP equal to the first n_{init} IDs in NDL
 19: **else:**
 20: set NIP equal to the IDs in NDL
 21: send an NIP

ALGORITHM 1: Initialization phase for anchor nodes.

the localization error, which represents the problem of error propagation in an iterative localization scenario. Therefore, an error control procedure must be implemented to improve the localization performance. In this paper, we use an error control method based on the neighboring geometric features

Input: The unique ID of current node ID_c
Define: HopInfo Packet $HIP = \{ID, position, hopInfo\}$
NeighborInfo Packet $NIP = \{ID_1, \dots, ID_{n_init}\}$
Anchor HopInfo List
 $AHL = [ID_p, position_p, hopInfo_p]$
Neighbor Distance List $NDL = [ID_p, d_i]$

Procedure:
1: **if** received an HIP :
2: parse the HIP to obtain $ID, position, hopInfo$
3: **if** ID in $AHL.ID$ and ID is an anchor:
4: **if** $hopInfo < AHL[ID].hopInfo$
5: $AHL[ID].hopInfo = hopInfo$
6: **else if** ID is an anchor:
7: append $[ID, position, hopInfo]$ to AHL
8: send an HIP
9: **if** $hopInfo == 1$:
10: record ID and RSSI
11: **if** received an NIP :
12: parse the NIP
13: **if** ID_c in NIP and $sizeof(AHL) > 2$:
14: **for** i from 1 to $sizeof(AHL)$:
15: estimate d_i to each anchor based on $hopInfo$ and the RSSI
16: estimate the position of ID_c based on position in AHL and a series of d_i
17: change node type to pseudo-anchor

ALGORITHM 2: Initialization phase for blind nodes.

of the target. The features used to estimate the localization confidence are discussed in Section 3.2.

Blind node j begins localization by broadcasting a packet containing a localization request and its unique ID. The neighbor of the target that received the request packet will return a location packet that contains the position of the node and the node type. The target collects all the location packets and the associated RSSI during the communication. Then, the target counts the number of anchor nodes in its neighborhood, which is denoted as $pLen$. If the value of $pLen$ is greater than or equal to 3 (which is the minimum requirement for multilateration), the target calculates its geometric features ($ALC_j, ALR_j,$ and ALS_j). Before calculating the features, the target should first estimate the distance to the anchors based on the RSSI values via (1). After determining the distances and positions of the anchor nodes, the geometric features of the current target node are estimated using (13), (14), and (15).

After discovering the current geometric features, a one-class support vector classifier is used for the localization decision. Conventional binary classification methods require both positive and negative training samples. However, guaranteeing that most of the negative states are included in the training is difficult. The OC-SVM differs from traditional classifiers in that only one class of training samples is required. The OC-SVM has demonstrated good performance for anomaly and fault detection.

By detecting anomalies during location estimation, we treat the error control problem as a one-class classification problem in which the OC-SVM is used with a linear kernel to detect good reference geometries and reject the others. The OC-SVM maps the input samples to a high-

dimensional feature space with a kernel function and then searches for a hyperplane to separate the mapping points from the origin (in this feature space) [24, 25]. We take the training set $D = \{p_i\}$, $p_i \in R^K$, AND $1 \leq i \leq n$ where p_i represents the i th training sample, K is the number of visual features, and n denotes the total number of samples. A kernel function ϕ is assumed, and it maps from the original space R^K to the infinite dimensional space χ , which satisfies $\phi(p_i) \in \chi$. Thus, the hyperplane is used to perform the following classification:

$$\begin{aligned} \min \quad & \frac{1}{2} \|\mathbf{w}\|^2 + \frac{1}{vn} \sum_{i=1}^n \xi_i - \rho, \\ \text{s.t.} \quad & \mathbf{w} \cdot \phi(p_i) \geq \rho - \xi_i, \xi_i \geq 0, \end{aligned} \quad (18)$$

where \mathbf{w} is the normal vector of the hyperplane, ρ represents the interval between the hyperplane and the original point, ξ_i is the slack variable corresponding to the i th sample that punishes the points that deviate from the hyperplane, and $v \in [0, 1]$ indicates the compromise between the maximum interval and the penalty term. A Lagrange function is used to derive the hyperplane, and the decision equation is obtained as follows:

$$f(p) = \text{sgn}((\mathbf{w} \cdot \phi(p)) - \rho). \quad (19)$$

If $f(p) \geq 0$, then p is assigned to the targets; otherwise, p is assigned to false alarms.

4. Localization Simulation

In this section, the localization algorithm described above is validated with a simulation experiment. First, we conduct a simulation in a grid sensor deployment that directly assesses the performance of the error control. Then, we further examine the algorithm performance in different environmental conditions for a randomly deployed sensor arrangement, including the number of anchor nodes, radio communication range, and RSSI variance. Finally, we compare the proposed algorithm to several global localization methods.

In the simulation, 400 sensor nodes are deployed in a 2D square environment with an edge length equal to 200 meters. The details of the simulation parameters are shown in Table 1. Not all the simulations in this section follow the value settings in the table. The differences will be described in the specific simulations.

4.1. Performance Metrics. The localization error in a 2D plane is defined as follows:

$$e = \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2}, \quad (20)$$

where (x_i, y_i) is the true ground position of the blind node, and (\hat{x}_i, \hat{y}_i) is the estimated position.

The relative localization error is defined as follows:

$$\bar{e} = \frac{1}{MR} \sum_{i=1}^M \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2}, \quad (21)$$

where M is the number of blind nodes and R is the radio communication range. This relative localization error is used to analyze the algorithm performance for different radio communication ranges R .

The Cramer-Rao lower bound derived in [26] is used to evaluate the algorithm. Based on the log-normal channel model, the Fisher information matrix is as follows:

$$F = \begin{bmatrix} F_{xx} & F_{xy} \\ F_{xy} & F_{yy} \end{bmatrix}, \quad (22)$$

and

$$\begin{aligned} F_{xx} &= b \sum_{i=1}^m \frac{(x - x_i)^2}{d_i^4}, \\ F_{yy} &= b \sum_{i=1}^m \frac{(y - y_i)^2}{d_i^4}, \\ F_{xy} &= b \sum_{i=1}^m \frac{(x - x_i)(y - y_i)}{d_i^4}, \\ b &= \left(\frac{10\eta}{\sigma \ln 10} \right)^2, \end{aligned} \quad (23)$$

where m is the number of anchor nodes. Then, the Cramer-Rao lower bound is given as follows:

$$\text{CRLB} = E[(x - \hat{x})^2 + (y - \hat{y})^2] \geq \frac{F_{xx} + F_{yy}}{F_{xx}F_{yy} - F_{xy}^2}. \quad (24)$$

4.2. Localization Results in Grid Deployment. Figure 5 shows a heat map of the direct localization results. In the figure, each square represents a sensor node, and the center of the square is the true ground position of the node. The square with a white background and a triangle in its center is an anchor node, and the other colored squares are blind nodes. Different colors are used to represent the localization error of each blind node, as indicated by the color bar on the right. Large localization errors are reflected by colors close to magenta, which represents an error of two meters, and the black-colored squares have a localization error larger than two meters. The two subfigures display the localization error when the number of anchor nodes is equal to 3 and 10. Additionally, in each anchor node scheme, we present four stages of localization results, in which the percentage of located blind nodes is 25%, 50%, 75%, and 100%.

As the figure shows, in the 100% stage, all the blind nodes in the network are located for each scheme, and the localization error is less than two meters, with the exception of a node in the corner of scheme (a). The localization starts from the neighborhood of the anchor nodes and then spreads all over the network. Certain blind nodes that have large localization error in the early stage exhibit a good localization result by the end of the localization. During the localization, a blind node with a bad localization result can be found and relocated with a better reference combination, thereby representing the effect of the proposed error control method. This method recognizes a bad localization condition based on the geometric features of the anchor node and improves the localization result by selecting a good reference combination.

4.3. Evaluation of Decision Parameter ν . The localization results shown in Section 4.2 can be further improved by tuning the localization parameter ν , as discussed in Section 3.3. Simulation experiments are conducted to determine the effect of parameter ν on the localization accuracy of the proposed algorithm. Figure 6 shows the localization error for parameter ν when the RSSI variance is $\sigma = 8$. Additionally, four different anchor node schemes denoted as NA (number of anchor nodes) are considered to examine parameter ν . Figure 6 shows that the error has a minimum value when ν varies from 0.1 to 1. When $\nu = 0$, the algorithm makes looser localization decisions, which means that more blind nodes will be located, even if the associated location confidence is bad. When $\nu = 1$, the algorithm makes strict decisions. Figure 6 shows that the algorithm obtains the minimum localization error when parameter ν is equal to 0.5 and that this value is stable for four different numbers of anchor nodes when the RSSI variance is equal to 8 dBm.

4.4. Comparison with the Nonerror Control Method. Figure 7 plots the cumulative distribution function of localization errors for different numbers of anchor nodes and shows that the proposed error control algorithm (the curve with upward triangular markers) can largely increase the localization accuracy compared with the iterative localization algorithm

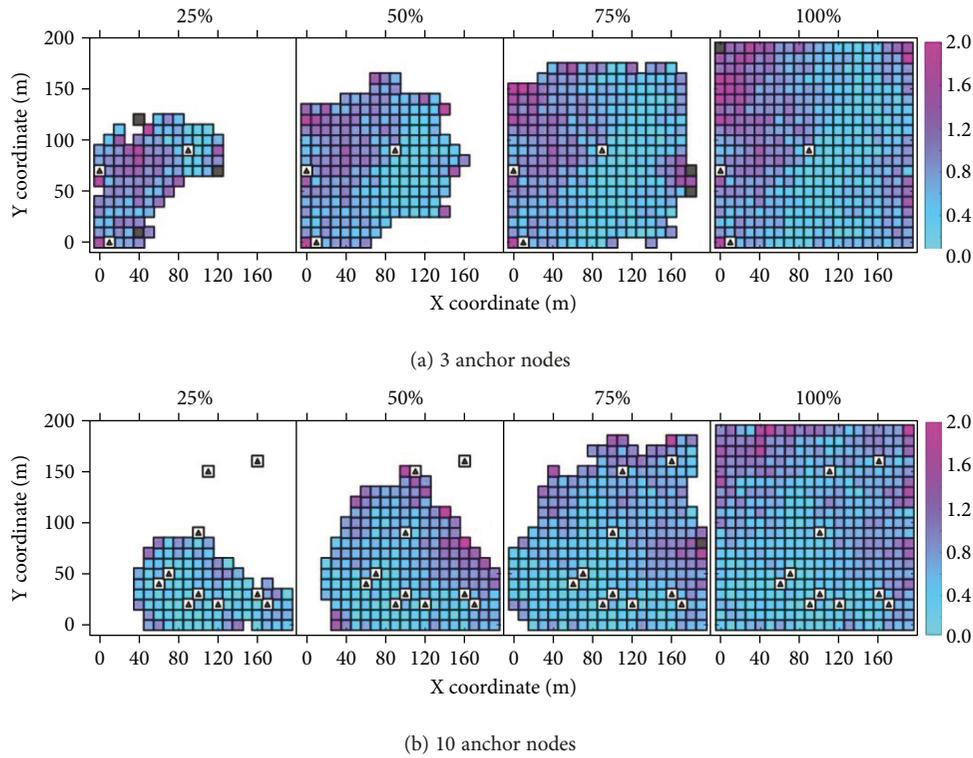


FIGURE 5: Localization error heat map with four different numbers of anchor nodes: (a) 3 anchor nodes are set and (b) 10 anchor nodes are set.

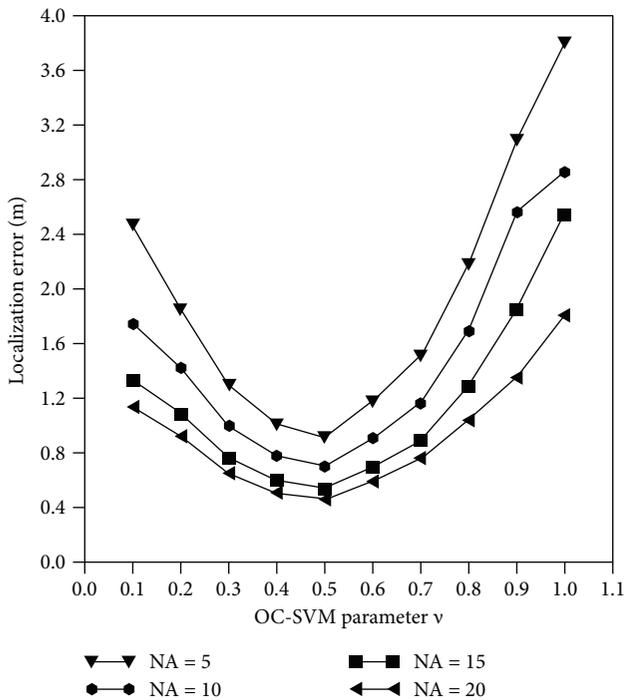


FIGURE 6: Performance of parameter ν for different numbers of anchor nodes when $\sigma = 8$.

without error control (the curve with dot markers). The value in the legend is the ratio of anchor nodes (1.25% represents 5 anchor nodes for 400 sensor nodes). The localization error decreases with an increase in the number of anchor nodes.

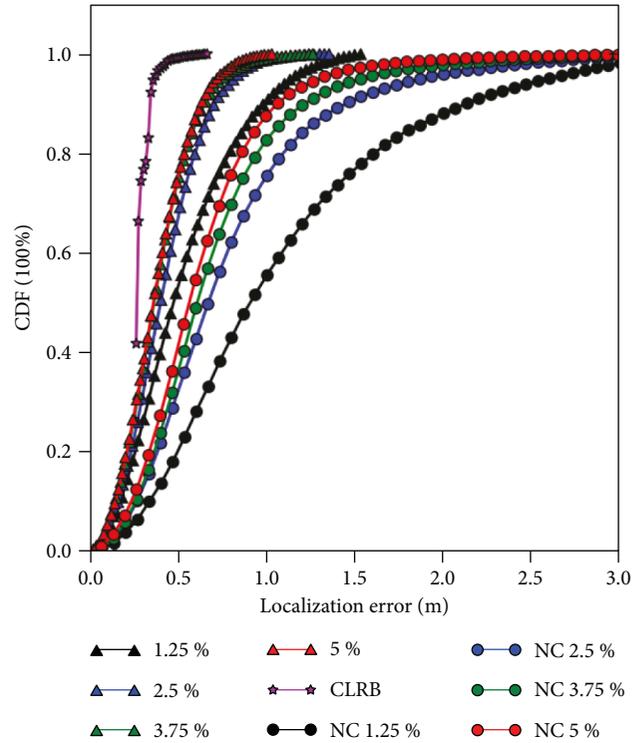


FIGURE 7: CDF for different numbers of anchor nodes when error control is used or not used.

However, when the ratio of anchor nodes is larger than 2.5% (the number of anchor nodes is larger than 10), the performance gradually improves as the number of anchor nodes

further increases. The green curve and the red curve with upward triangular markers are almost identical. Additionally, the localization results are similar to those of the CRLB (calculated by (24)) when the ratio of anchor nodes is equal to 5% (20 anchor nodes).

4.5. Evaluation of the RSSI Variance σ . Equation (8) displays that the variances in localization errors are caused by the variances in the ranging method. The ranging method used in the proposed algorithm is RSSI based. Therefore, the RSSI variances affect the uncertainty of the localization results. Simulation experiments are conducted, and the results are shown in Figure 8. We set the RSSI variances to range from 1 dBm to 10 dBm and set the means of the RSSI errors to zero. We also vary the number of anchor nodes (indicated by NA) and the radio communication range (indicated by R). The results show that the accuracy of the algorithm gradually decreases as the value of σ increases. Increases in the environmental variance will cause variance in the RSSI signal, produce noise in the ranging results, and affect the localization results. When the number of anchor nodes (NA = 20) and the radio communication range (R = 50 m) are large, the algorithm is more stable relative to cases with fewer anchor nodes and a shorter radio range. More anchor nodes provide more localization references, which means that the proposed algorithm has a greater chance of selecting a better localization condition to minimize the localization error. A longer radio communication range can similarly increase the number of anchor nodes and improve the localization performance.

4.6. Comparison with Global Localization Methods. Finally, we compare the proposed algorithm with several global localization algorithms for varying numbers of anchor nodes. These localization algorithms are MDS [27], SDP [28], and SPA [29], which are global in nature, although heuristic algorithms have been used to calculate the associated distributions. Among these algorithms, SPA is simple and easy to implement in real sensor networks; thus, it is used as a baseline comparison. To compare the performances, 400 sensor nodes are randomly deployed in a 200 m \times 200 m square area. The number of anchor nodes is varied from 3 to 10. The RSSI variance is set to 2 dBm, and the mean of the RSSI error is set to zero. The decision parameter ν of our algorithm is set to 0.5. The values of the other simulation parameters are the same as those in Table 1. The results of the simulation are shown in Figure 9. The localization error obviously decreases with an increase in the number of anchor nodes. When the number of anchor nodes is small, the DSP and SPA algorithms perform poorly, and the MDS algorithm performs relatively well. Figure 10 plots the accuracy of each localization method when the number of anchor nodes is set to 10. The proposed algorithm has the best localization accuracy compared with the other three algorithms.

5. Performance Results for Experimental Data

In this section, we conduct some experimental analyses to evaluate the efficiency of the proposed localization algorithm based on a real WSN deployed in an outdoor football field.

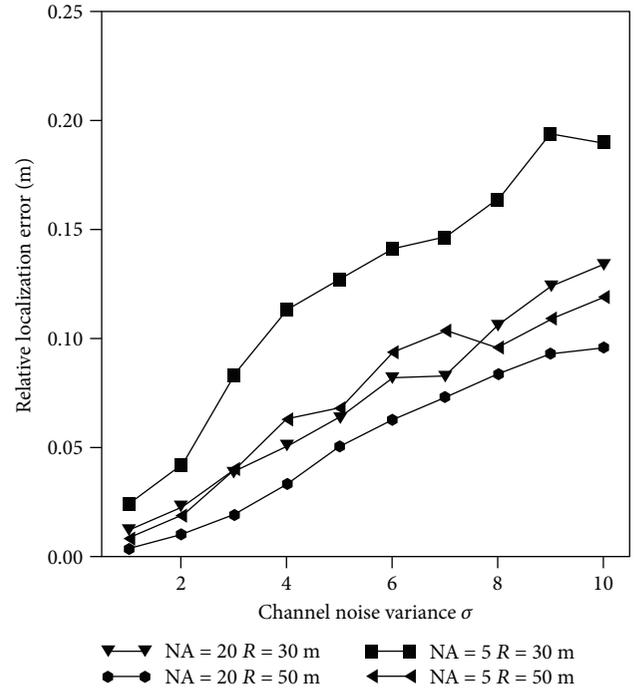


FIGURE 8: Performance of RSSI variance σ for different numbers of anchor nodes and radio ranges.

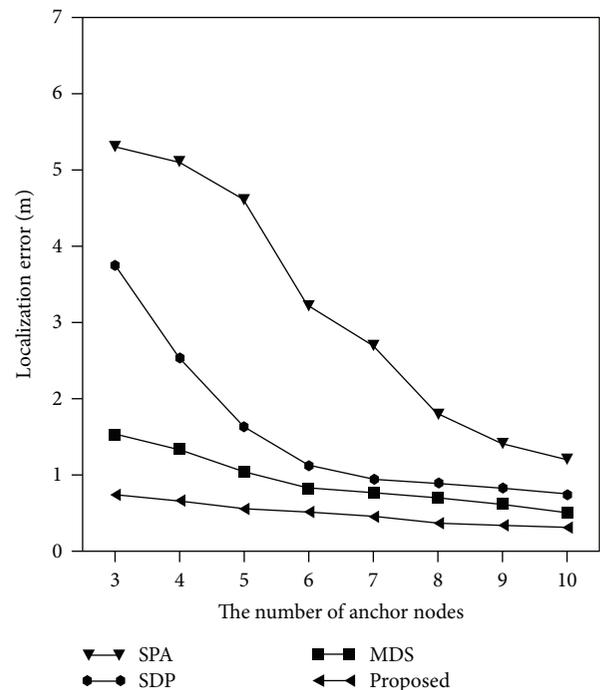


FIGURE 9: Comparison of the performance of different algorithms with different numbers of anchor nodes.

There are two stages in the experiment. In the first stage, we deploy a sensor network with only one blind node to test the efficiency of the weighted least squares position estimator. In the second stage, a multihop sensor network is

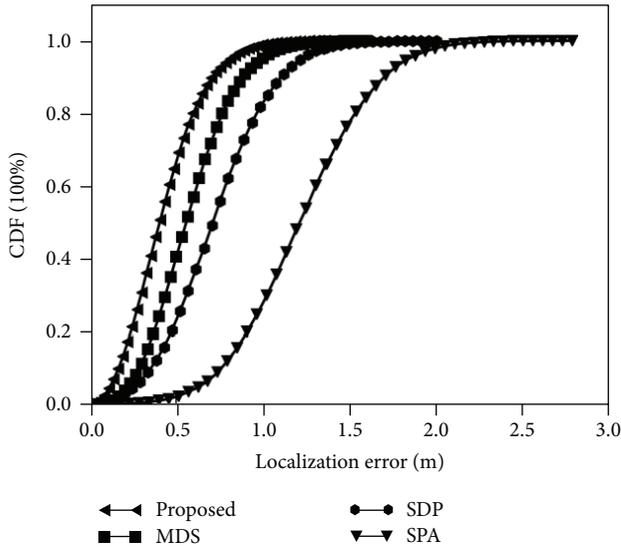


FIGURE 10: Accuracy of the different localization algorithms.

deployed to test the efficiency of the proposed localization algorithm with error control.

The sensor nodes are based on a TI CC2650 wireless radio transceiver, which runs a TI-RTOS real-time embedded system and is powered by a 4800 mAh battery bank. A 2.4 GHz whip antenna is attached to the sensor nodes, and the nodes are synchronized by an FTSP time synchronization algorithm.

5.1. Position Estimation Experiment. In this experiment, four sensor nodes working as anchor nodes are deployed at the vertexes of a square, which are indicated by the blue circles in Figure 11. The edge of the square is 10 meters, and 16 positions are selected as the test points in the square, which are indicated by the red dots. The test points are set in a grid with a 2-meter edge length, as the figure shows. Additionally, a blind node is established among the test points. There is also a sink node, which is set 10 meters from the test area; a laptop is connected to the sink node by a USB cable to store the data collected. During the experiments, the anchor nodes will send a message to each blind node. After receiving the message from the reference nodes, the blind node extracts the RSS value of the communication and stores the data in a local buffer for further use. When 10 RSS values from each anchor node are collected, the blind node estimates its position using linear least squares (LS) and weighted least squares (WLS) methods. Then, the RSS data and the position estimation results are uploaded to the sink node and stored in the laptop. On the laptop, a nonlinear position estimation method is implemented to estimate the position of the blind node, and the results are printed with the results from the blind node. For each test point, the procedure described above is repeated 50 times; then, the radio parameters are changed, and the experiment is repeated. After all the radio parameters have been tested, the blind node will be moved to the next position.

Figure 12 shows the localization error for each test point with the different position estimation methods: WLS, LS, and

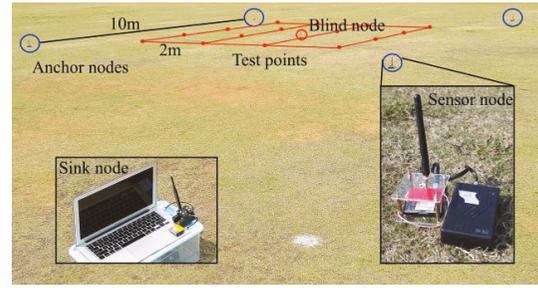


FIGURE 11: Deployment area in the position estimation experiment.

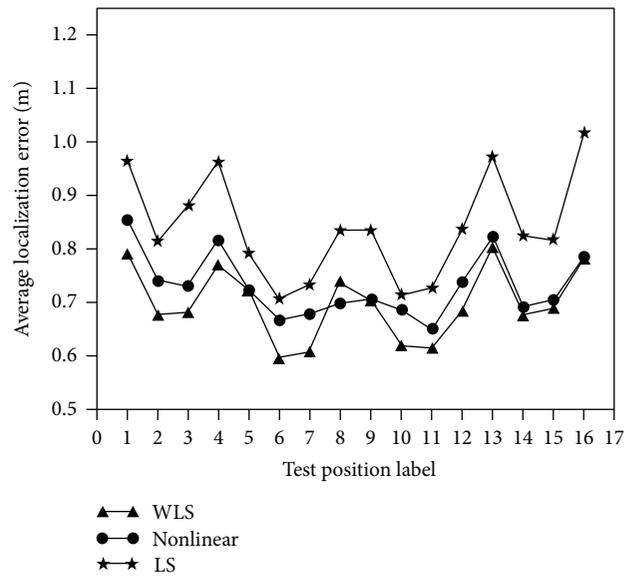


FIGURE 12: Average localization error for each test point with different estimation methods.

nonlinear methods. The WLS position estimator exhibits the lowest error for all 16 test points (the average localization errors are 0.68 m for WLS, 0.72 m for nonlinear, and 0.85 m for LS); furthermore, the variance of the error is only 0.1 m. Moreover, the errors for different groups of test points display different trends. For test points 6, 7, 10, and 11, the localization errors are similar, with average errors of 0.63 m, 0.65 m, and 0.75 m for the WLS, nonlinear, and LS methods, respectively. These localization errors are lower than the average level for different methods, as these four test points are in the center of the test area, and the distances between the blind node and each anchor node are similar, which reflects a stable localization pattern, as discussed in Section 3.1. For test points 2, 3, 5, 8, 9, 12, 14, and 15, the average localization errors are 0.68 m, 0.72 m, and 0.84 m for the WLS, nonlinear, and LS methods, respectively. These errors are higher than the errors in the first group noted above. For test points 1, 4, 13, and 16, the average localization errors are 0.78 m, 0.82 m, and 0.97 m for the WLS, nonlinear, and LS methods, respectively, and these values are much higher than those of the other two groups, as these test points reflect an unstable localization pattern.

Figures 13 and 14 show the average localization error in the experiment with different values of the log-normal channel model parameter. The WLS algorithm exhibits better localization accuracy than the LS algorithm. Additionally, the accuracy of the WLS method is better than that of the nonlinear method for most of the channel model parameters. Considering the high complexity of the nonlinear method, WLS displays good efficiency in localization procedures. The curves in the figures exhibit a valley, which reflects the optimal parameters for the current channel environment that minimizes the localization errors. In the current situation, the approximate values of the parameters are $\eta = 3.7$ and $A = -34$ dBm. In real applications of a WSN, the wireless radio channel model should not be constant, and WLS should exhibit higher robustness to channel model variance, which can help improve the localization accuracy of the localization algorithm.

5.2. Localization Experiment with Error Control. In the second experiment, we deploy 16 sensor nodes in a 4×4 square grid area, as shown in Figure 15. The positions of the nodes are indicated by red dots in the subfigure in Figure 15, and the vertical or horizontal distance between neighboring nodes is 10 meters. To construct a multihop network in a relatively small area, we must modify the traditional sensor network. During the experiments, the transmission power of the sensor nodes is set lower than that in a traditional network, and a constant neighborhood is written into the nodes, which means that node 1 can only be heard by sensors 2, 5, and 6 (sensor 6 can only be heard by sensors 1, 2, 3, 5, 7, 9, 10, and 11). To test the efficiency of the iterative localization algorithm for different anchor setting conditions, we conduct the experiment following an experimental scheme, as shown in Table 5. With different numbers of anchor nodes and different anchor node setups, the experiment is separated into 10 different schemes. In the first five schemes, three sensor nodes are selected as the anchor node. Not all the possible combinations of the anchor nodes are covered in these five schemes, as the layout of the network is symmetrical. In the last five schemes, the same considerations are applied.

During the experiments, the sensor nodes run the localization algorithm introduced in Section 3. Every time the blind node obtains a position estimation, it uploads the position results and the RSS value used for localization to the sink node. Additionally, the data received by the sink node are transmitted to the laptop through a USB cable. When the position of the blind node becomes stable, the experiment is stopped. For each experimental scheme, we run the experiment 50 times; then, we change the radio channel parameters and perform the experiment again. After all the radio parameters are assessed, the nodes are established for the next experimental scheme.

The distance estimation accuracy is important for the proposed localization algorithm. For the given deployments of the sensor network, the real distances between sensor nodes are 10 m, 14 m, 20 m, 22 m, 28 m, 30 m, 32 m, 36 m, and 42 m. During the experiment, the distance estimation results of the sensor nodes are collected, and the estimation errors are shown in Figure 16. The average distance estimation error

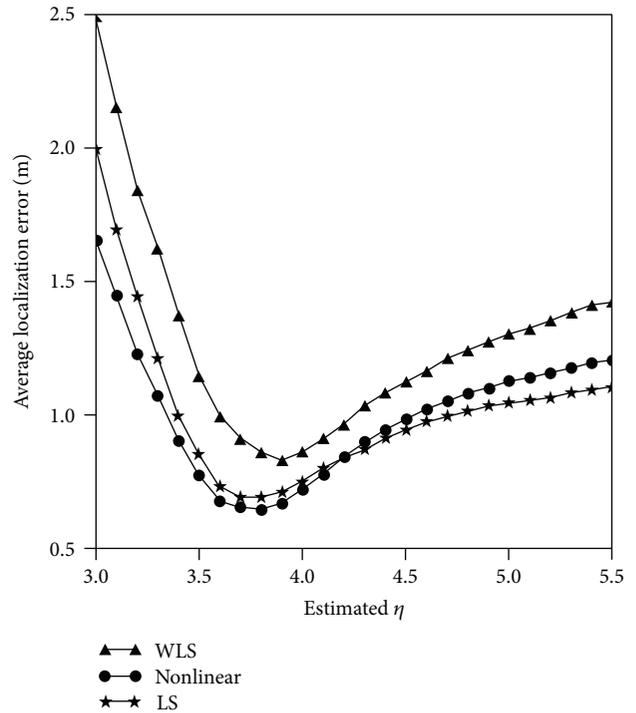


FIGURE 13: Average localization error for different values of η and $A = -34$ dBm.

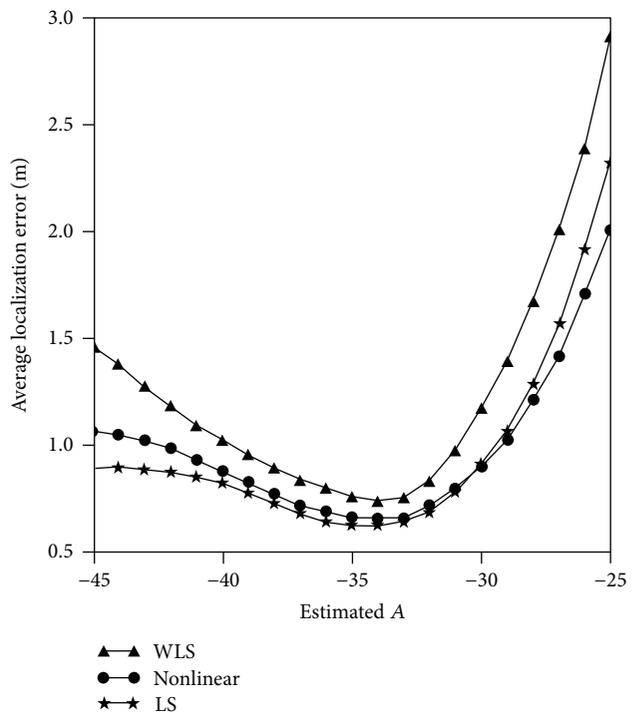


FIGURE 14: Average localization error for different values of A and $\eta = 3.7$.

exhibits little difference as the real distance increases from 10 m to 42 m. However, the variance in the error considerably increases as the real distance increases.

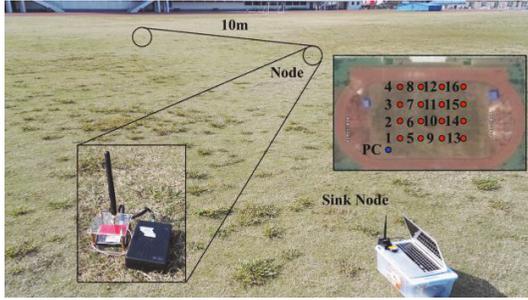


FIGURE 15: Deployment area in the iterative localization experiment.

TABLE 5: Experimental schemes.

Scheme	Anchor ID	Anchor ID	Anchor ID	Anchor ID
T1	1	2	5	×
T2	1	2	3	×
T3	1	4	13	×
T4	1	10	16	×
T5	7	10	11	×
F1	1	2	5	6
F2	1	4	13	16
F3	1	2	3	4
F4	1	7	10	16
F5	6	7	10	11

The localization errors of all the schemes listed in Table 5 are shown in Table 6. The results show that the average localization error for all the schemes is 1.23 m and that the variance of the error is 0.8m. Moreover, a comparison of the results of the schemes with three anchor nodes (series T) with those of the schemes with four anchor nodes (series F) reveals that the average localization errors are similar. This finding suggests that the proposed localization algorithm displays similar performance for sensor networks with different anchor node settings because the error control mechanism of the proposed localization algorithm always tries to find a better localization reference to update the position estimation. The results of series F exhibit slightly less error variance because these schemes have one more anchor node, which can increase the localization stability. The errors of schemes T2 and T4 are higher than those of the other results in the T series, as the layout of the anchor nodes in these two schemes tends to be linear. Although the error control mechanism can reduce the localization error to 2.42 m (T4 1.9m), the initial localization instability will affect the final results. Comparing schemes T3 and T5, the average error of T5 is lower than that of T3 because the anchor nodes in T3 are set in a corner. The anchor node layout of T3 can cause the blind nodes on the edge of the area to have a higher localization error because an unstable localization pattern exists. Additionally, the proposed algorithm is sensitive to the initial positions of the anchor nodes. However, in real applications of large-scale sensor node deployment, the network is always fully random with more

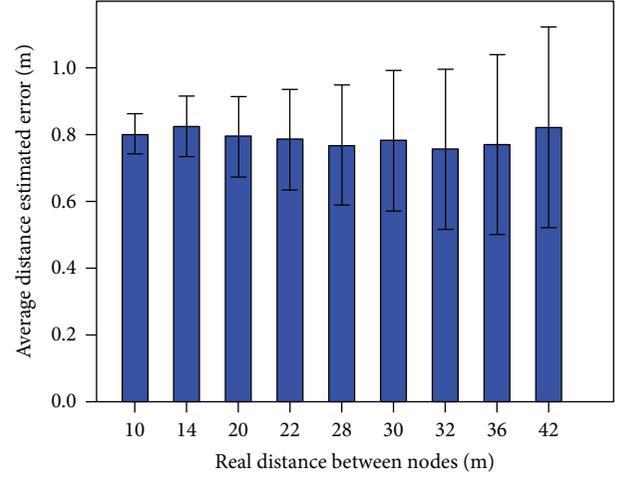


FIGURE 16: Average distance estimation error.

TABLE 6: Localization results for different schemes.

Scheme	Average error	Error variance
T1	0.62	0.38
T2	2.31	1.41
T3	1.43	0.6
T4	1.83	1.24
T5	0.67	0.43
F1	0.61	0.36
F2	0.47	0.35
F3	2.38	1.28
F4	1.71	1.11
F5	0.8	0.41
Total	0.8	0.41

than three anchor nodes; therefore, this situation would rarely occur.

5.3. Computational Complexity. In some applications of WSNs, the computational ability is constrained, which requires the localization to be performed on chip. Computational complexity, which is defined as the number of operations performed by the algorithm, is used to evaluate the possibility of chip implementation.

The proposed localization algorithm has three parts, which are presented in Algorithms 1, 2, and 3. Algorithms 1 and 2 are the initialization phase of the localization algorithm. Additionally, the computational complexity is $O(n^2)$ for a numerical sorting algorithm and a linear LS algorithm. n is the number of anchor nodes used for position estimation. Algorithm 3 is more complex; it involves feature computation, OC-SVM prediction, and WLS position estimation. Thus, the computational complexity is $O(3n^3)$. This computational complexity is not high in real applications, as the value of n is never high (the average number of anchor nodes used by a blind node to perform localization is 9 in our simulation). An anchor node selection scheme can also be

Input: The unique ID of current node ID_c
 The node Type of current node $Type_c$

Define: Localization Request Packet LRP
 Location Packet $LP = \{ID, Position, Type\}$
 Reference List $RL = [ID_p, p_p, d_p]$
 Localization Confidence Score LCS
 Estimated Position EP
 Confidence List $CL = [ID, EP, LCS, Type]$
 Localization Decision LD

Procedures:

- 1: send an LRP
- 2: **If** received an LP :
- 3: parse the LP to obtain $ID, position, Type$
- 4: obtain the $RSSI$
- 5: **if** $Type == reference$:
- 6: obtain distance estimation d via Eq. (1)
- 7: append $ID, position, d$, to RL
- 8: **If** $sizeof(RL) > 3$:
- 9: calculate current geometric features via Eqs. (13) (14) and (15)
- 10: calculate the LCS of this localization by Eq. (19)
- 11: **If** $LCS > LCS$ in the last CL
- 12: input the features to one-class SVM classifier to obtain a decision result LD
- 13: **If** $LD == localization$:
- 14: estimate the location of current node ID_c by Eq. (10).
- 15: change the type of current node to pseudo-anchor
- 16: append $[ID_c, EP, LCS, pseudo-anchor]$ to CL

ALGORITHM 3: Error control phase for blind nodes.

designed into the algorithm to further decrease the computational complexity of the algorithm.

6. Conclusions

This paper presents a novel iterative localization algorithm that mitigates the effects of error propagation with an anchor geometric feature-based error control mechanism. The advantage of this approach is that it can be used when only a few anchor nodes are involved. Traditional methods often require a large number of well-placed anchor nodes to work well; this limitation is overcome by using a combination of DV-hop localization and a novel error control mechanism. First, we analyze the geometric relationships between the blind node and the anchor nodes. Several features are proposed to describe the geometric features. Then, an OC-SVM-based decision machine is trained to select a good localization condition for the blind node. The performance of the algorithm is assessed by simulation experiments as well as real-life outdoor experiments, and the results show that the localization accuracy is improved compared with that of several global localization methods.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (11402121).

References

- [1] M. Obaidat and S. Misra, *Principles of Wireless Sensor Networks*, Cambridge University Press, 2014.
- [2] Y. Yang and C. Wang, *Wireless Rechargeable Sensor Networks*, Springer International Publishing, 2015.
- [3] M. Demirbas, Xuming Lu, and P. Singla, "An in-network querying framework for wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 8, pp. 1202–1215, 2009.
- [4] V. K. Chaurasiya, N. Jain, and G. C. Nandi, "A novel distance estimation approach for 3D localization in wireless sensor network using multi dimensional scaling," *Information Fusion*, vol. 15, pp. 5–18, 2014.
- [5] J. K. Lee, Y. Kim, J. H. Lee, and S. C. Kim, "An efficient three-dimensional localization scheme using trilateration in wireless sensor networks," *IEEE Communications Letters*, vol. 18, no. 9, pp. 1591–1594, 2014.
- [6] W. D. Wang and Q. X. Zhu, "RSS-based Monte Carlo localisation for mobile sensor networks," *IET Communications*, vol. 2, no. 5, pp. 673–681, 2008.
- [7] D. Wang, T. Pei, Y. Zhou, and Y. Wei, "Distributed multi-object localisation by consensus on compressive sampling

- received signal strength fingerprints,” *IET Communications*, vol. 9, no. 14, pp. 1738–1745, 2015.
- [8] F. Yin, C. Fritsche, F. Gustafsson, and A. M. Zoubir, “TOA-based robust wireless geolocation and Cramér-Rao lower bound analysis in harsh LOS/NLOS environments,” *IEEE Transactions on Signal Processing*, vol. 61, no. 9, pp. 2243–2255, 2013.
- [9] N. Wu, Y. Xiong, H. Wang, and J. Kuang, “A performance limit of TOA-based location-aware wireless networks with ranging outliers,” *IEEE Communications Letters*, vol. 19, no. 8, pp. 1414–1417, 2015.
- [10] J. Xu, C. L. Law, and M. Ma, “Performance of time-difference-of-arrival ultra wideband indoor localisation,” *IET Science, Measurement & Technology*, vol. 5, no. 2, pp. 46–53, 2011.
- [11] K. C. Ho, “Bias reduction for an explicit solution of source localization using TDOA,” *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2101–2114, 2012.
- [12] Z. Irahauten, H. Nikoogar, and M. Klepper, “A joint ToA/DoA technique for 2D/3D UWB localization in indoor multipath environment,” in *2012 IEEE International Conference on Communications (ICC)*, pp. 4499–4503, Ottawa, ON, Canada, June 2012.
- [13] H. Chenji and R. Stoleru, “Toward accurate mobile sensor network localization in noisy environments,” *IEEE Transactions on Mobile Computing*, vol. 12, no. 6, pp. 1094–1106, 2013.
- [14] C. Medina, J. Segura, and Á. de la Torre, “Ultrasound indoor positioning system based on a low-power wireless sensor network providing sub-centimeter accuracy,” *Sensors*, vol. 13, no. 3, pp. 3501–3526, 2013.
- [15] J. Liu, Y. Zhang, and F. Zhao, “Robust distributed node localization with error management,” in *2006 International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pp. 250–261, Florence, Italy, May 2006.
- [16] M. Erol-Kantarci, H. T. Mouftah, and S. Oktug, “A survey of architectures and localization techniques for underwater acoustic sensor networks,” *IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 487–502, 2011.
- [17] H. Suo, J. Wan, L. Huang, and C. Zou, “Issues and challenges of wireless sensor networks localization in emerging applications,” in *2012 International Conference on Computer Science and Electronics Engineering*, pp. 447–451, Hangzhou, China, March 2012.
- [18] J. Liu and Y. Zhang, “Error control in distributed node self-localization,” *EURASIP Journal on Advances in Signal Processing*, vol. 2008, no. 1, Article ID 162587, 2007.
- [19] Z. Yang and Y. Liu, “Quality of trilateration: confidence-based iterative localization,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 5, pp. 631–640, 2010.
- [20] X. Wu, S. Tan, and Y. He, “Effective error control of iterative localization for wireless sensor networks,” *AEU - International Journal of Electronics and Communications*, vol. 67, no. 5, pp. 397–405, 2013.
- [21] Y. Hu, L. Zhang, L. Gao, X. Ma, and E. Ding, “Linear system construction of multilateration based on error propagation estimation,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, 2016.
- [22] T. K. Sarkar, Zhong Ji, Kyungjung Kim, A. Medouri, and M. Salazar-Palma, “A survey of various propagation models for mobile communication,” *IEEE Antennas and Propagation Magazine*, vol. 45, no. 3, pp. 51–82, 2003.
- [23] D. Niculescu and B. Nath, “DV based positioning in ad hoc networks,” *Telecommunication Systems*, vol. 22, no. 1–4, pp. 267–280, 2003.
- [24] Y. Chen, X. S. Zhou, and T. S. Huang, “One-class SVM for learning in image retrieval,” in *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, pp. 34–37, Thessaloniki, Greece, 2001.
- [25] Y. Guerbai, Y. Chibani, and B. Hadjadji, “The effective use of the one-class SVM classifier for handwritten signature verification based on writer-independent parameters,” *Pattern Recognition*, vol. 48, no. 1, pp. 103–113, 2015.
- [26] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O’Dea, “Relative location estimation in wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137–2148, 2003.
- [27] Y. Shang, W. Rumi, Y. Zhang, and M. Fromherz, “Localization from connectivity in sensor networks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 15, no. 11, pp. 961–974, 2004.
- [28] P. Biswas and Y. Ye, “Semidefinite programming for ad hoc wireless sensor network localization,” in *Proceedings of the third international symposium on Information processing in sensor networks - IPSN’04*, pp. 46–54, Berkeley, CA, USA, April 2004.
- [29] K. Whitehouse and X. Jiang, “Calamari: a localization system for sensor networks,” 2004, <http://www.cs.virginia.edu/~whitehouse/research/localization/>.

