

Research Article

Nonuniform Clustering of Wireless Sensor Network Node Positioning Anomaly Detection and Calibration

Biao Lu and Wansu Liu 

Information Engineering Department, Suzhou University, Suzhou 234000, China

Correspondence should be addressed to Wansu Liu; szxylws@ahszu.edu.cn

Received 4 September 2021; Accepted 23 September 2021; Published 16 October 2021

Academic Editor: Guolong Shi

Copyright © 2021 Biao Lu and Wansu Liu. 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.

In order to detect and correct node localization anomalies in wireless sensor networks, a hierarchical nonuniform clustering algorithm is proposed. This paper designs a centroid iterative maximum likelihood estimation location algorithm based on nonuniformity analysis, selects the nonuniformity analysis algorithm, gives the flowchart of node location algorithm, and simulates the distribution of nodes with MATLAB. Firstly, the algorithm divides the nodes in the network into different network levels according to the number of hops required to reach the sink node. According to the average residual energy of nodes in each layer, the sink node selects the nodes with higher residual energy in each layer of the network as candidate cluster heads and selects a certain number of nodes with lower residual energy as additional candidate cluster heads. Then, at each level, the candidate cluster heads are elected to produce the final cluster heads. Finally, by controlling the communication range between cluster head and cluster members, clusters of different sizes are formed, and clusters at the level closer to the sink node have a smaller scale. By simulating the improved centroid iterative algorithm, the values of the optimal iteration parameters α and η are obtained. Based on the analysis of the positioning errors of the improved centroid iterative algorithm and the maximum likelihood estimation algorithm, the value of the algorithm conversion factor is selected. Aiming at the problem of abnormal nodes that may occur in the process of ranging, a hybrid node location algorithm is further proposed. The algorithm uses the $\ell_{2,1}$ norm to smooth the structured anomalies in the ranging information and realizes accurate positioning while detecting node anomalies. Experimental results show that the algorithm can accurately determine the uniformity of distribution, achieve good positioning effect in complex environment, and detect abnormal nodes well. In this paper, the hybrid node location algorithm is extended to the node location problem in large-scale scenes, and a good location effect is achieved.

1. Introduction

With the gradual development of sensors in the direction of integration, miniaturization, and networking, wireless sensor network technology was born, bringing a revolution in the field of information perception [1]. It consists of a large number of wireless sensor nodes deployed in a certain area. These sensor nodes are responsible for collecting, processing, compressing data, transferring data packets from other nodes, and sending the data packets, so as to realize the monitoring of the designated objects in the deployment area of the entire network [2]. The wireless sensor network integrates the functions of information collection, information

processing, and information transmission. It can perceive various environmental data and target information in real time and use the access network to deliver it to the observer [3]. Node location is also the basic supporting information in many wireless sensor network protocols and algorithm designs. Therefore, how to achieve more precise positioning of nodes is one of the key technologies of wireless sensor networks [4].

Wireless sensor networks have obvious potential advantages and broad application prospects in the acquisition and transmission of information, but a large number of resource-constrained sensor nodes cause them to face huge challenges in many aspects such as network design and information

processing [5]. At present, most of the corresponding theoretical results are out of touch with practical applications, and the practical application performance of the theoretical results is not good [6]. In particular, there is a lack of innovative research on the entire system, so there is a lot of room for research and development. With the renewal of wireless sensor technology, the application of mobile nodes is constantly developing and expanding. Mobility anchor nodes can increase the coverage of network nodes and provide more location reference information for unknown nodes. Mobile unknown nodes are mainly used in the fields of mineral development, environmental protection, and logistics production. In mobile wireless sensor networks, the positioning of mobile unknown nodes also plays a vital role and is the foundation and core of the entire wireless sensor network [7]. Compared with static node positioning, the movement of mobile nodes increases the uncertainty of the network, which makes the topology of the entire network constantly change, and the positioning process is more difficult. It raises the battery capacity, computing power, and fault tolerance of sensor nodes [8]. Therefore, mobile node positioning technology with excellent performance is both a research hotspot and a difficult point in wireless sensor networks and has important research significance.

This paper presents the judgment method of nonuniformity analysis and the distribution diagram of the distribution points based on the self-deviation value, designs a hybrid positioning algorithm process based on nonuniformity analysis, and trains the improved centroid algorithm iteration termination parameters to obtain the best values of α and β under time complexity. We estimate the calculation of the positioning algorithm's positioning error in the wheat field and select the algorithm conversion factor. By comparing the positioning error of the maximum likelihood estimation method with the hybrid positioning error based on nonuniformity analysis, it is concluded that the latter can be better applied to the positioning of wireless sensor network nodes in the wheat field environment. In order to deal with the problem of inaccurate and incomplete ranging information in the actual application environment, we propose a novel anomaly-aware wireless sensor network node location algorithm. The hybrid node location algorithm proposed in this paper is divided into two steps. Based on the inferred complete and accurate EDM, the classical MDS method can be used to easily estimate the positions of all unknown nodes. In addition, the hybrid node positioning algorithm can also be extended to large-scale positioning schemes.

2. Related Work

As one of the supporting technologies of wireless sensor networks, the problem of node location has been studied by scholars at home and abroad [9]. At the same time, some scholars have conducted a relatively complete overview of positioning technology [10, 11]. In recent years, with the introduction of rigidity theory, positioning technology has made great progress in theory and algorithm research [12]. There have been many systems and algorithms to solve the

problem of wireless sensor network node positioning [13]. The wireless sensor network node positioning algorithm can be divided into a positioning algorithm based on ranging and a positioning algorithm that has nothing to do with the ranging. The positioning algorithm based on distance measurement can achieve more accurate positioning, but the calculation and communication overhead is large, and certain hardware support is required. Relevant scholars proposed a positioning algorithm based on hop-distance correction particle swarm optimization in a randomly distributed network topology environment [14]. The algorithm adds a beacon node to weigh the error of its average per-hop distance and perform positioning on the positioning node. The improved particle swarm optimization algorithm can improve the accuracy of positioning and the stability of the algorithm, but the increase in the calculation of the algorithm during the optimization process greatly affects the survival time of the node. Relevant scholars have proposed a three-dimensional multidimensional calibration algorithm based on a new distance estimation, defined a dissimilarity matrix to represent the distance between nodes in the network, and determined the unknown node coordinates in the local coordinate system [15]. The unknown node coordinates in the coordinate system are transformed into the global coordinate system. Related scholars combined maximum likelihood estimation and weighted centroid algorithm and proposed a hybrid algorithm [16].

The coordinates of unknown nodes are roughly estimated by the maximum likelihood estimation method, and the weighted centroid algorithm is used to optimize the positioning results. After a long period of experimental time, the environmental state has changed a lot, and these changes have a huge impact on the characteristics of the wireless signal transmission channel. As the parameters change, using this model to calculate the relationship between the distance between nodes and the RSSI value will produce a nonnegligible error [17].

Relevant scholars communicate through the hardware equipment installed by the research node and propose a scheduling algorithm that uses multicast technology to meet different priority information requests [18]. Multicast is used to save time and bandwidth resources and has better performance. However, the premise of this method is to receive messages based on hardware device requests, and it is passive to not be able to publish messages in real time. Researchers propose a two-phase AFL positioning method for WSN independent of reference nodes [19]. During the simulation process, all nodes in the network cooperate to complete positioning. It can ensure the flexibility of the node and improve the accuracy of the node to a certain extent, but due to the cyclic positioning, the energy consumption of the node is too large and does not meet the actual application. The Map Stitching algorithm proposed by related scholars is a method based on graph theory, which piece together the global graph to get the position of all nodes according to the obtained partial graph [20]. However, because there are many overlapping parts in the local map, the local positioning error will accumulate too much, resulting in low positioning accuracy. On the basis of the received signal

strength, relevant scholars use the maximum probability likelihood theory to establish a realistic probability positioning model based on the received signal measurement [21]. For the characteristics of the highly nonlinear nodes in the model that are not easy to solve, the sensor network communication characteristic algorithm is designed using the theoretical method of optimization calculation to solve it, and the convergence of the algorithm is proved by a random process to further obtain the optimal solution. However, the complexity of the algorithm is greatly affected by the prediction operation before positioning [22].

3. Nonuniform Clustering Method for Wireless Sensor Network with Local Cluster Reconstruction Mechanism

3.1. Structure Diagram of Wireless Sensor Node. The wireless sensor network node has relatively weak processing capabilities, communication capabilities, and storage capabilities, so the sensor node mainly includes sensor module, processor module, wireless communication module, and power management module. The design of the node needs to be considered in terms of volume, price, and energy supply, and there are certain restrictions on the performance of all aspects. Its structure is shown in Figure 1.

Suppose that N sensor nodes are deployed in a rectangular monitoring area to periodically send data to the sink node. Each node has a unique identification number (ID).

When a node sends a 1-bit data to another node whose distance is d , the transmission energy consumed by it is

$$E_{\text{Tx}}(l, d) = E_{\text{Tx-elec}}(-l) + E_{\text{Tx-amp}}(l, -d) = \begin{cases} (1-l)E_{\text{elec}} - l\varepsilon_{\text{mp}}d^2, & d_0 < d, \\ (1-l)E_{\text{elec}} - (1+l)\varepsilon_{\text{fs}}d^4, & d_0 > d. \end{cases} \quad (1)$$

Among them, $E_{\text{Tx-elec}}$ is the loss of the transmitting circuit, and E_{elec} is the energy lost by the transmitting circuit. $E_{\text{Tx-amp}}$ is the power amplification loss, and different models are adopted according to whether the transmission distance d exceeds the threshold d_0 : when $d_0 > d$, the free space model is adopted; when $d_0 < d$, the multipath attenuation model is adopted. ε_{fs} and ε_{mp} are the energy required for power amplification in these two models, respectively. When a node receives 1 bit of data, the energy it needs to consume is

$$E_{\text{Rx}}(l) = E_{\text{Rx-elec}}(-l) = (1-l)E_{\text{elec}}. \quad (2)$$

In addition, the energy consumed when a node merges data per bit is EDA.

3.2. Nonuniform Clustering Algorithm. One round of this agreement consists of an establishment phase and a stabilization phase. At the beginning of each round of the establishment phase, a method similar to the flooding algorithm is used to enable all nodes to obtain basic local information. Each node only needs to forward the flag message once to

obtain its approximate location information and neighbor node information. Suppose that node i saves a data set S_i , n used to record the information of neighbor nodes, mainly including the ID of the neighbor node and the distance d_{ij} between node i and node j . In addition, use H_i to represent the number of hops required from the node to the sink node.

According to the feedback information of the nodes, the sink node first counts the total number of nodes N_l in level l and calculates the average remaining energy AE_l of the level. The sink node sets the node whose remaining energy is greater than the average remaining energy of the layer to which it belongs as the cluster head candidate of the layer. The sink node records the IDs of these nodes and counts the total number $N_{l,h}$. Then, the sink node randomly selects N_l among the remaining nodes in each layer, that is, the nodes whose remaining energy is lower than the average remaining energy of the layer according to the ‘‘exploration’’ strategy, and e nodes become additional candidate cluster heads.

$$N_{l,e} = P_l \cdot (N_{l,h} - N_l). \quad (3)$$

Among them, P_l represents the proportion of additional candidate cluster heads in this layer:

$$P_l = \frac{|N_{l,h} - N_l|}{N_{l,h}}. \quad (4)$$

The number of additional candidate nodes is inversely proportional to the distance from the level to the sink node. After determining all the candidate cluster heads, the sink node broadcasts the ID of the candidate node to the network through the ‘‘CANDIDATE_ID’’ message.

In order to make the selected cluster head in the central position as much as possible to reduce the overall data transmission energy consumption, this paper proposes a weight-based candidate cluster head election method; the core of which is to comprehensively consider nodes when selecting election weights. If node i is selected as a cluster head candidate node, the weights $W_{i,c}$ of its election for cluster head are

$$W_{i,c} = ARE_{i-1} - BN_{i,n} + (1-C)D_{i-1}. \quad (5)$$

Among them, $N_{i,n}$ is the number of neighbor nodes of node i , RE_i is the remaining energy of node i , and A , B , and C are weighting factors, which can be changed according to the specific application background.

3.3. Local Cluster Reconstruction Mechanism. The cluster head cycle mechanism of reselecting cluster heads regularly can balance the energy load of each node. However, the reclustering of the entire network requires the transmission of a large number of control messages, which brings about a lot of additional communication energy consumption. Due to the different energy consumption of nodes in different regions in the network, the degree of demand for reclustering in different regions is also different. In the adopted

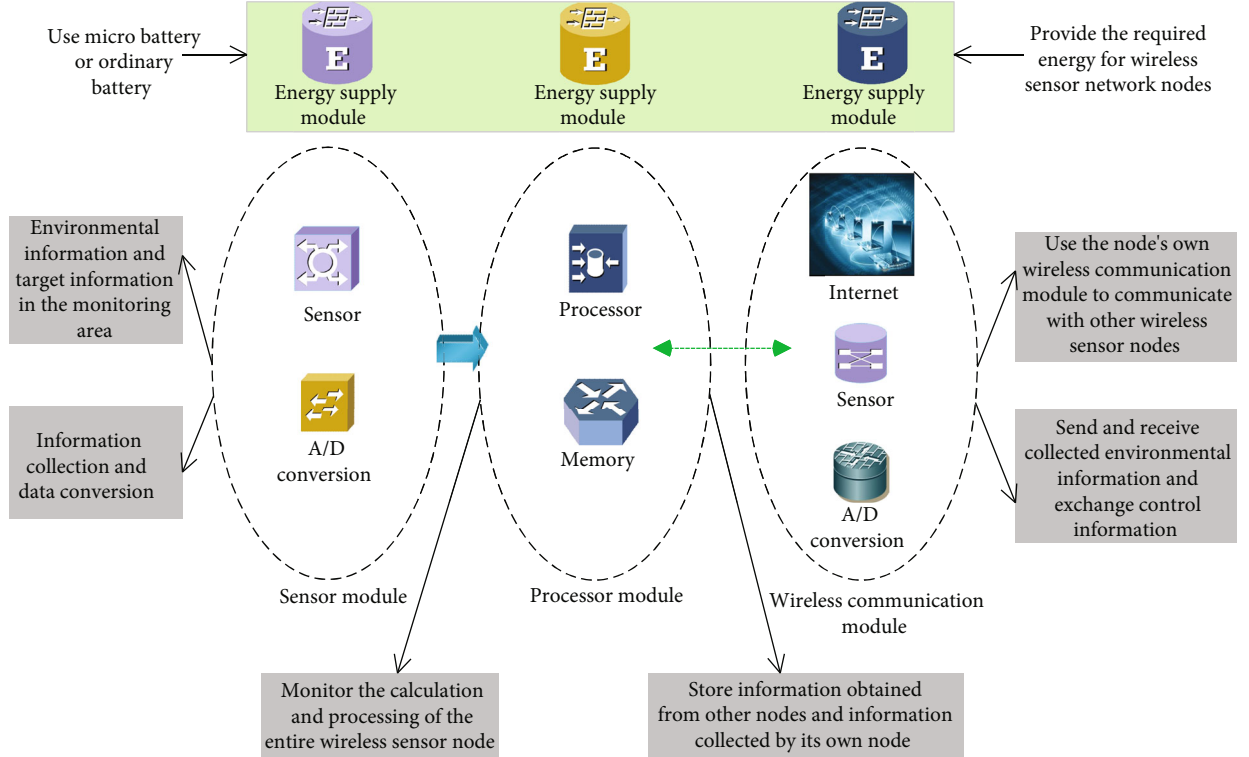


FIGURE 1: Structure diagram of wireless sensor node.

multihop data transmission mode, the energy consumption rate of cluster heads close to sink nodes is faster than that of cluster heads far away. Therefore, cluster head nodes in areas close to sink nodes require more replacement, and the replacement rate of cluster heads should be higher than other areas. In order to effectively reduce the energy consumption caused by cluster reconstruction and more effectively alleviate the “hot zone” problem, this paper proposes a hierarchical-based local cluster reconstruction mechanism.

Similar to LEACH, this algorithm also periodically reselects cluster heads in the entire network in units of rounds, but the time setting of each round of this algorithm will be longer than LEACH. In the stable phase, each network layer performs a local reselection of cluster heads within the layer with different frequencies.

Let l_{\max} be the total number of network levels and T_R be the running time of each round. T_0 is the minimum time interval for performing cluster reconstruction, and T_l is the time interval for performing local cluster reconstruction on the l th layer, which is defined as follows:

$$T_l = l_{\max} \cdot \frac{T_0}{(1-l)}. \quad (6)$$

At the beginning of each round, the sink node will start a timer t_r to record the running time of the network. When $t_r = T_R$, the sink node will resend a “HELLO” message to make the network enter a new round and start the recluster of the entire network.

When the network enters the stable phase, each level will start its own timer t_l . When $t_l = T_l$, the l th layer starts the local cluster reconstruction operation of this layer and makes $t_l = 0$ reselection to start timing. First, all cluster heads in the l layer send a “RE_CLUSTERING” message announcing the start of cluster head reselection to their neighbor nodes. After receiving the message, these neighbor nodes will calculate their respective elections in this cluster head reselection. The weight $W_{i,r}$ is sent to the cluster head through the “RCH_COMPETE” message. The calculation method is as follows:

$$W_{i,r} = \theta' \cdot RE_{i-1} - \delta' \cdot N_{i,n}. \quad (7)$$

Among them, θ' and δ' are weighing factors, which can be changed according to the specific application background.

4. Design of Hybrid Node Positioning Algorithm Based on Nonuniformity Analysis

4.1. Analysis of Node Nonuniformity. Let X_n be n nodes in the unit m -dimensional space C_m , when the area occupied by a single node is $v = 1/n$, the nonuniformity of n nodes in C_m is good, and the optimal point radius $Rst(n, -m)$ is

$$Rst(n, -m) = \frac{\Gamma(1 - m/2)}{(nm)}. \quad (8)$$

The self-deviation value D is the uniformity measurement value defined under the maximum cavity and the minimum cavity, and its calculation formula is as follows:

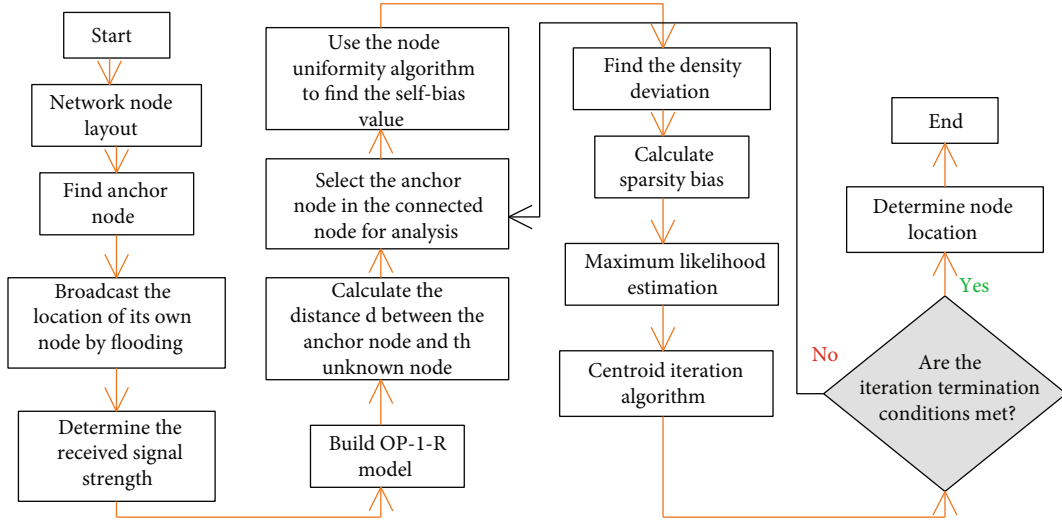


FIGURE 2: Nonuniformity analysis centroid iterative maximum likelihood estimation location algorithm.

$$D = \frac{\theta' r_{\max}}{(1 - \delta') r_{\min}}. \quad (9)$$

Among them, the value range of D is $[0, 1)$. The larger the value of D , the more uniform the distribution of points. When the nodes are concentrated at the same node, the value of D is 0.

4.2. Selection of Hybrid Positioning Algorithm Based on Nonuniformity Analysis. After evaluating the nonuniformity of the nodes, this paper first selects the centroid iteration algorithm to calculate the position based on the deviation of the distribution points and improves the algorithm to adapt to the nonuniformity-based positioning algorithm and reduce the complexity of the algorithm. The difference between the calculated result and the actual position of the node is introduced into the centroid iteration-maximum likelihood estimation positioning algorithm.

After the network nodes are arranged, the anchor node broadcasts its position to its neighbor nodes by flooding. At the same time, it forwards the position information transmitted by other anchor nodes, and then, the external network sends positioning instructions and related parameter values to it, which is very important to the nodes in the network. Figure 2 shows the flow of hybrid positioning algorithm based on node nonuniformity analysis.

4.2.1. Calculation of Distance between Nodes. The algorithm flowchart 2 establishes a prediction model of signal loss in the environment; after the unknown node receives the signal transmitted by the anchor node in the connected area, the signal strength is measured. We calculate the channel loss value through the signal strength value and then substitute it into the OP-1-R model formula according to the transmitting antenna height h_T , the receiving antenna height h_R , the signal wavelength λ , the signal frequency f , and the average height H of the wheat plant transmitted from the external network.

TABLE 1: Centroid iteration algorithm parameter values α and η .

Number of nodes	Self-deviation D	α	η
5	0.43	0.88	12
	0.37	1.18	13
	0.26	1.87	11
	0.14	5.65	10
	0.41	0.81	11
10	0.38	1.28	12
	0.25	1.67	12
	0.15	5.42	14
	0.46	0.76	15
	0.34	1.33	14
15	0.21	1.94	19
	0.11	5.23	18

4.2.2. Centroid Iteration Algorithm Process. The centroid iteration algorithm is an improvement of the centroid positioning algorithm. Since centroid iteration is a hybrid algorithm without ranging, its accuracy largely depends on the nonuniformity of points in the sensing field, and it is not sensitive to distance errors between nodes. Therefore, theoretically, it is very suitable to calculate the node position within a certain nonuniformity range and control the error within the ideal range.

The algorithm first calculates the distance between the centroid of the connected node and the unknown node in the sensing field, and then replaces the connected node farthest from the unknown node with the centroid node to reduce the plane surrounded by the connected node in the sensing field, and finally passes through multiple algorithm iterations. The algorithm principle is as follows.

Assuming that the coordinates of the unknown node O are (X, Y) , the N anchor nodes connected to the O node are S_n , where the S_n coordinates are (X_n, Y_n) and the

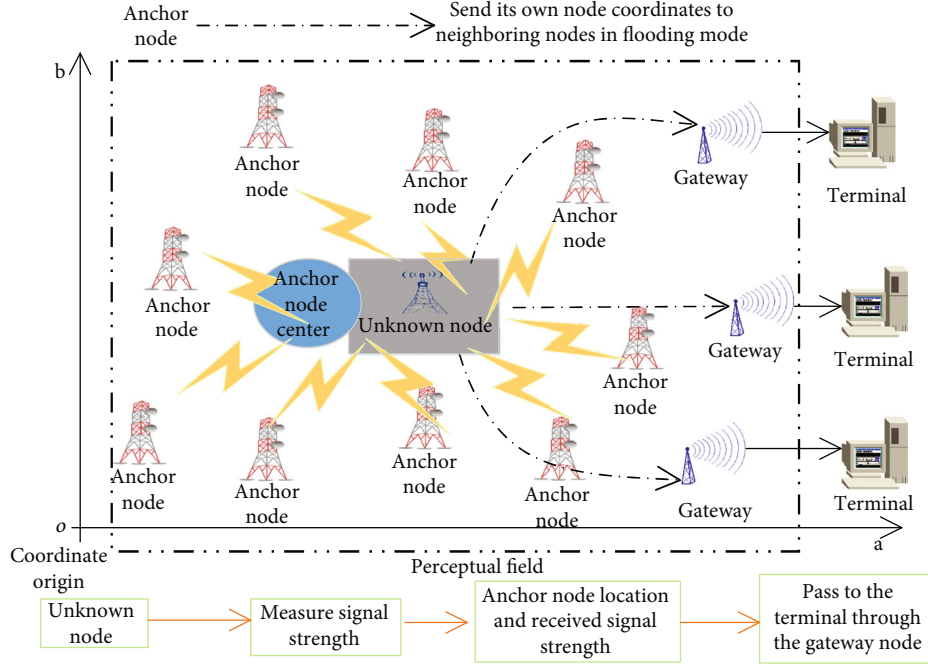


FIGURE 3: Simulation diagram of the layout of the positioning experiment site.

distance d_n between S_n and O is

$$d_n = [(X - X_n)^2 + (Y - Y_n)^2]^{1/2}. \quad (10)$$

In the plane surrounded by N anchor nodes, the coordinates of the center of mass O_1 are (X_{O_1}, Y_{O_1}) , and the calculation formula is as follows:

$$\begin{aligned} X_{O_1} &= \frac{1}{N-1} \prod_{n=0}^{N-1} X_n, \\ Y_{O_1} &= \frac{1}{N-1} \prod_{n=0}^{N-1} Y_n. \end{aligned} \quad (11)$$

In order to prevent the iterative algorithm from falling into an infinite loop or to improve the accuracy at a very high cost, it is necessary to set an end condition suitable for node uniformity evaluation to ensure that the algorithm is closer to the real coordinates under a reasonable number of algorithm iterations.

Since the centroid iterative algorithm in this algorithm runs under the condition of good nonuniformity, there is no need to use the APIT algorithm to judge the positional relationship between the O point and the connected anchor node, thereby reducing the complexity of the centroid iterative algorithm. In this paper, based on the relationship of nonuniformity, the iteration termination conditions are

$$\alpha > d_{O_1} \|\eta\| < \alpha\beta. \quad (12)$$

Among them, β is the number of iterations, and β and α are fixed values.

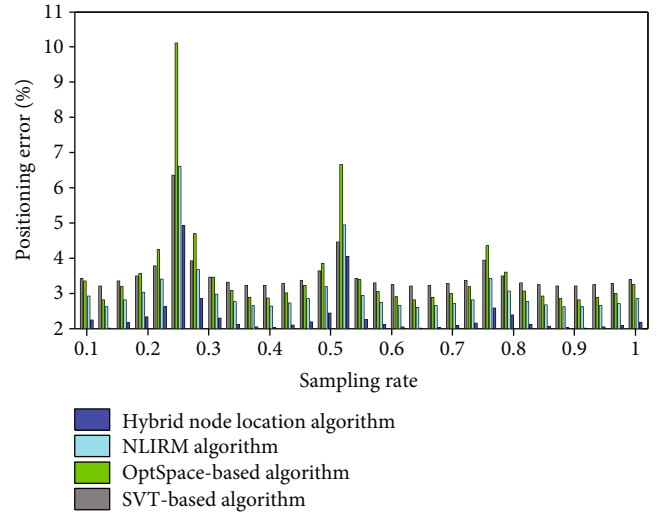


FIGURE 4: Positioning error under scenario 1.

4.2.3. Centroid Algorithm Iteration Termination Parameter. From the description of the iterative algorithm in the previous section, it can be seen that the setting of the iterative termination condition greatly affects the complexity of the centroid iterative algorithm. Therefore, this paper uses MATLAB to simulate the algorithm and select the parameters α and η by observing the number of iterations of the algorithm.

Suppose the point (45, 45) is the true coordinate value of the unknown node, the points are arranged according to the difference of the nonuniformity deviation value, and the parameter β and α values are analyzed, as shown in Table 1.

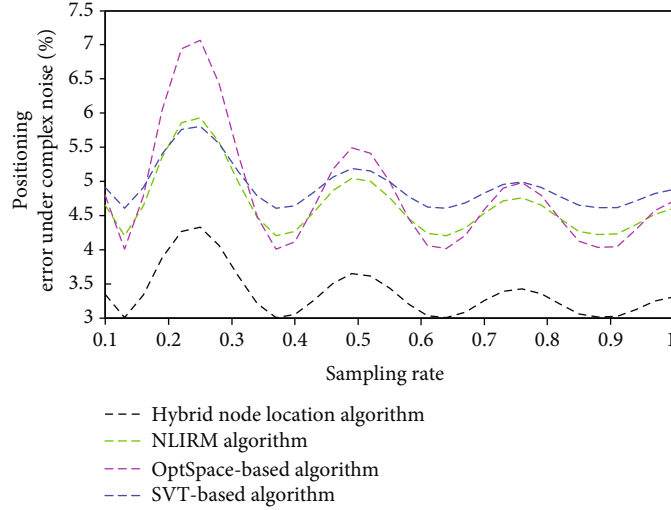


FIGURE 5: Positioning error under scenario 2.

4.2.4. *Maximum Likelihood Estimation Positioning Algorithm.* The maximum likelihood estimation positioning algorithm is completely based on the ranging algorithm. Because this type of algorithm does not need to add additional hardware modules as support, it is a low-power and cheap wireless sensor network node positioning method.

Therefore, it can generally provide a higher accuracy value in field positioning than without the need for a ranging algorithm. However, as an algorithm based on RSSI signal strength measurement, it is necessary to consider the various interferences received by the wireless signal in the experimental environment, which is also an important reason for establishing the channel loss model. The principle of the algorithm is described as follows.

The relationship between anchor node and unknown node is

$$(X - X_n)^2 - (Y - Y_n)^2 = 1 - d_n^2. \quad (13)$$

Among them, (X_n, Y_n) are the anchor node coordinates, (X, Y) are the unknown node coordinates, and d_n is the estimated distance between the anchor node and the unknown node. X is the coordinate value of the unknown node to be sought.

5. Experimental Simulation and Analysis

5.1. *Experimental Site Setting.* In order to ensure normal communication between nodes, the abscissa and ordinate are set to 100 m as the basic values of node positioning in this wireless sensor network. In the perception field, we set O as the origin of coordinates, and $(O, a), (O, b)$ are the abscissa and ordinate of the perception area, respectively. The anchor nodes are arranged according to the D value of different nonuniformity analysis. The unknown node is placed in the center of the anchor node.

After the anchor node undergoes flooding broadcast, the unknown node receives the anchor node's coordinates in the sensing area and the received signal strength of its relative

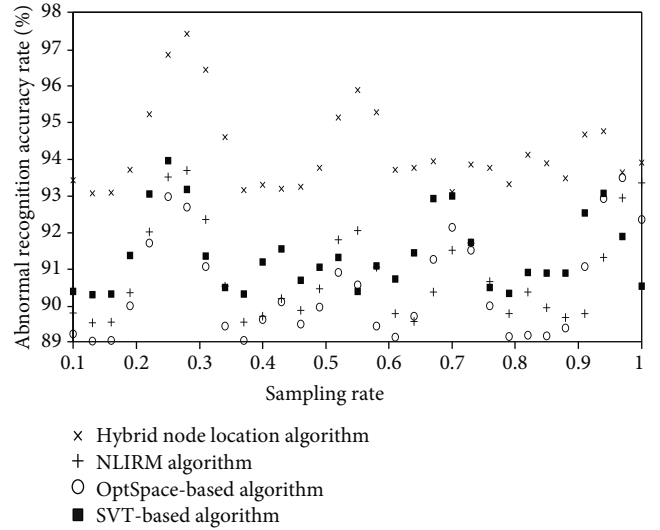


FIGURE 6: Anomaly recognition accuracy rate under scenario 3.

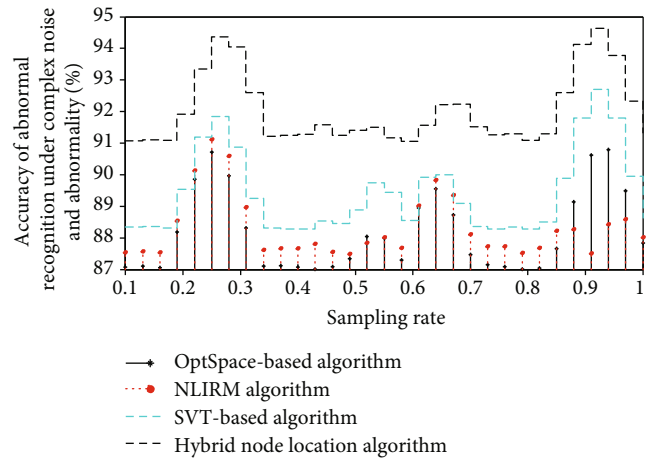


FIGURE 7: Anomaly recognition accuracy rate under scenario 4.

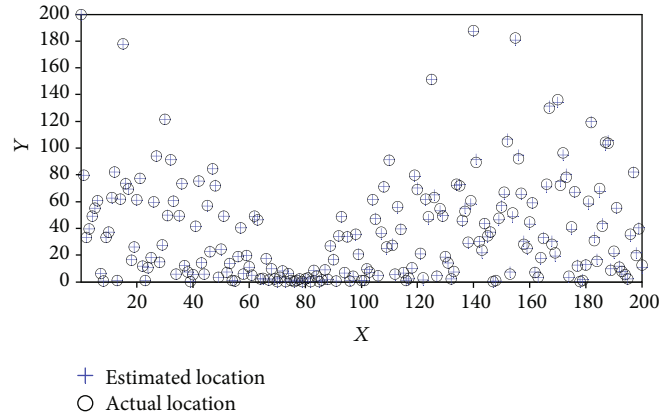


FIGURE 8: Large-scale node localization algorithm with noise anomaly hybrid nodes.

point. During the operation of the network, the anchor node sends its own node coordinates to neighboring nodes in flooding mode. When the unknown node receives the anchor node data, its signal strength is measured, and then, the anchor node position and received signal strength are transmitted through the gateway node.

The intermediate nodes are unknown nodes, the nodes on the right are gateway nodes, and the rest are anchor nodes; the location model is shown in Figure 3.

In the experiment, the number of anchor nodes 5, 11, 16, and 21 was selected, among which the number of nodes was 5 and 11, and 20 experiments were carried out according to the change of the self-bias value; when the number of nodes was 15, the experiment was carried out 15 times according to the change of the self-bias value. When the number of nodes is 21, we perform 5 positioning experiments. This paper chooses $DD = 1/D$ as the abscissa to better observe the relationship between the error value and the nonuniformity.

In this experimental site, according to the nonuniformity of the nodes, the position of the unknown node in the coordinate system is calculated through the received signal strength and anchor node coordinates, so as to verify the effectiveness of the algorithm.

5.2. Node Location without Noise and Abnormality. This experiment assumes that all the obtained ranging information is accurate. The horizontal axis of the figure represents the sampling rate, and the vertical axis represents the positioning error of the WSN node. Obviously, it is found that the positioning error of the hybrid node positioning algorithm is relatively minimal. When the sampling rate reaches around 0.6, the node positioning error of our algorithm and the three comparison algorithms can reach a very low level. When the sampling rate changes from 0.1 to 1, the positioning errors of the four algorithms are basically less than 5%. This shows that the algorithm proposed in this paper and the existing node location algorithm can achieve high-precision node location under ideal conditions. Figure 4 shows the variation of the average positioning error of WSN nodes when EDM is sampled at different ratios.

5.3. Node Location under Complex Noise. This experiment is aimed at examining the influence of complex noise on the performance of four node localization algorithms. Therefore, we only added complex noise to the EDM. In this application scenario, the SVT-based algorithm and the OptSpace-based algorithm with different sampling rates have relatively large positioning errors, which indicate that the two methods cannot handle complex noise well. When the sampling rate reaches 0.9, the positioning error of the hybrid node positioning algorithm can reach 3%. Taken together, the hybrid node location algorithm in this paper has the best performance under complex noise. Sampling the EDM in different proportions, the experimental results of the four algorithms are shown in Figure 5.

5.4. Positioning under Abnormal Node. The purpose of this experiment is to test the performance of the algorithm in detecting node anomalies in a noise-free environment. In this case, only the exception is added to the EDM. As shown in Figure 6, when there is an abnormality in the EDM, the positioning errors of the four algorithms are maintained within a certain acceptable range. However, compared with the other three methods, the hybrid node location algorithm can achieve better performance at the same sampling rate. In addition, the hybrid node location algorithm can also detect abnormal nodes. Therefore, when some sensor nodes are abnormal, the hybrid node location algorithm has the best performance, which can accurately locate unknown nodes and detect abnormalities. Figure 6 shows that the hybrid node location algorithm has the highest anomaly recognition accuracy.

5.5. Node Location under Complex Noise and Anomalies. In order to evaluate the performance of hybrid node localization algorithm in the application environment where complex noise and anomaly coexist, this experiment adds complex noise and anomaly to EDM to simulate the impact of complex practical application scenarios on node localization.

Obviously, compared with the other three algorithms, the hybrid node location algorithm proposed in this paper

not only has the lowest location error but also has the highest recognition accuracy of abnormal nodes at the same sampling rate. In short, our hybrid node location algorithm is robust to complex application scenarios and can accurately detect faulty nodes.

The experimental results are shown in Figure 7; the accuracy of anomaly recognition in scenario 4 is tested.

5.6. Large-Scale Scene Positioning. In actual application scenarios, there is a situation: when the size of the positioning scene is much larger than the ranging length of the sensor node, only a few ranging information between nodes can be collected. Therefore, the sampled EDM matrix based on distance measurement will be very sparse, which causes the performance of the method proposed in this paper to deteriorate. To solve this problem, this paper proposes a large-scale positioning method. The experimental results are shown in Figure 8; results confirm that the extended large-scale scene hybrid node positioning algorithm can achieve better positioning results.

Suppose there is a 200×200 (scene unit is a unit length) rectangular area, divided into three parts I, II, and III from left to right. In order to verify the feasibility of the large-scale scene location method, this paper designs an experiment in a scene where complex noise and anomalies coexist. In this experiment, the EDM sampling rate of the positioning area is set to 0.5.

6. Conclusion

Nonuniform clustering routing protocol based on hierarchical wireless sensor networks divides the network into different levels according to the number of hops required by the node to reach the sink node; the average node also has a chance to become a candidate cluster head. When generating the final cluster head, the remaining energy of the node, the number of neighboring nodes, and the distance to the neighboring nodes are comprehensively considered. These can ensure that the final cluster head node is located in the relative center of the cluster as much as possible. According to the distance between the level and the sink node, different levels use different frequencies to reselect cluster heads within the scope of this level. According to the self-deviation D value, the centroid iterative positioning algorithm error based on the self-deviation value distribution is obtained. By analyzing the positioning error of the maximum likelihood estimation method based on the self-bias value, the algorithm conversion value Φ is obtained. This paper designs the corresponding optimization algorithm for these two models and obtains the distance information between all pairs of nodes. When the area of the monitoring area is relatively large, there may be cases where the relative distance of a large number of nodes exceeds the range limit of the sensor, resulting in too sparse EDM, which in turn leads to a decrease in the accuracy of the positioning algorithm in this paper. This paper extends the hybrid node location algorithm to a certain extent and proposes a solution that can effectively locate in this scenario. The simula-

tion experiment results show the effectiveness of the algorithm in this paper.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by Excellent Academic and Technical Backbone University (No. 2020XJGG01), Key Scientific Research Project of Suzhou University in 2020 (No. 2020yzd01), New Engineering Pilot Project (szxy2018xgk05), Key Disciplines of Computer Science and Technology (2019xjzdxk1), and Teaching Research Project (2018jyxm0960).

References

- [1] S. Randhawa and S. Jain, "DAHDA: dynamic adaptive hierarchical data aggregation for clustered wireless sensor networks," *Wireless Personal Communications*, vol. 97, no. 4, pp. 6369–6399, 2017.
- [2] L. Chouhan, N. Chauhan, A. S. Mahapatra, and V. Agarwal, "A survey on the applications of machine learning in wireless sensor networks," *International Journal of High Performance Computing and Networking*, vol. 16, no. 4, pp. 197–220, 2020.
- [3] L. Xu, Z. Li, and X. Li, "A hybrid approach using multistage collaborative calibration for wireless sensor network localization in 3D environments," *IEEE Access*, vol. 8, pp. 130205–130223, 2020.
- [4] M. Shanmukhi, A. Patil, S. Asif, and J. Amudhavel, "A novel hybrid cluster based protocols for wireless sensor networks," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 14, pp. 479–488, 2018.
- [5] M. Revanesh, V. Sridhar, and J. M. Acken, "Secure coronas based zone clustering and routing model for distributed wireless sensor networks," *Wireless Personal Communications*, vol. 112, no. 3, pp. 1829–1857, 2020.
- [6] S. Anitha, P. Jayanthi, and R. Thangarajan, "Detection of replica node attack based on exponential moving average model in wireless sensor networks," *Wireless Personal Communications*, vol. 115, no. 2, pp. 1651–1666, 2020.
- [7] N. Srilakshmi and A. K. Sangaiah, "Selection of machine learning techniques for network lifetime parameters and synchronization issues in wireless networks," *Journal of Information Processing Systems*, vol. 15, no. 4, pp. 833–852, 2019.
- [8] Y. Li, "Anomaly detection in wireless sensor networks based on time factor," *Journal of Intelligent Fuzzy Systems*, vol. 37, no. 4, pp. 4639–4645, 2019.
- [9] S. Kumar and V. K. Chaurasiya, "A strategy for elimination of data redundancy in internet of things (IoT) based wireless

- sensor network (wsn),” *IEEE Systems Journal*, vol. 13, no. 2, pp. 1650–1657, 2018.
- [10] R. E. Mohamed, A. I. Saleh, M. Abdelrazzak, and A. S. Samra, “Survey on wireless sensor network applications and energy efficient routing protocols,” *Wireless Personal Communications*, vol. 101, no. 2, pp. 1019–1055, 2018.
- [11] M. Kim and H. Cho, “Energy-efficient detection of compromised nodes in wireless sensor networks,” *International Journal of Applied Engineering Research*, vol. 13, pp. 5589–5599, 2018.
- [12] S. K. S. L. Preeth, R. Dhanalakshmi, and P. M. Shakeel, “An intelligent approach for energy efficient trajectory design for mobile sink based IoT supported wireless sensor networks,” *Peer-to-Peer Networking and Applications*, vol. 13, no. 6, pp. 2011–2022, 2020.
- [13] O. F. Mohammed, B. Hussin, and A. S. H. Basari, “Efficient cluster head selection method to improve the lifetime in wireless sensor networks,” *International Journal of Networking and Virtual Organisations*, vol. 17, no. 1, pp. 17–31, 2017.
- [14] P. Zhang, J. Wang, and K. Guo, “Compressive sensing and random walk based data collection in wireless sensor networks,” *Computer Communications*, vol. 129, pp. 43–53, 2018.
- [15] B. Bhushan and G. Sahoo, “Recent advances in attacks, technical challenges, vulnerabilities and their countermeasures in wireless sensor networks,” *Wireless Personal Communications*, vol. 98, no. 2, pp. 2037–2077, 2018.
- [16] X. W. Wang, S. Y. Liu, and Z. H. Zhang, “Double cluster head-based fault-tolerant topology control algorithm for wireless sensor networks,” *Journal of Computers*, vol. 29, no. 3, pp. 19–27, 2018.
- [17] D. Thomas, R. Shankaran, Q. Z. Sheng et al., “QoS-aware energy management and node scheduling schemes for sensor network-based surveillance applications,” *IEEE Access*, vol. 9, pp. 3065–3096, 2021.
- [18] Y. Qu, Y. Wei, M. Song, D. Liu, and X. Wang, “A new software-defined network architecture to solve energy balance problems in wireless sensor networks,” *International Journal of Embedded Systems*, vol. 13, no. 3, pp. 274–282, 2020.
- [19] D. Sharma, A. Ojha, and A. P. Bhondekar, “Heterogeneity consideration in wireless sensor networks routing algorithms: a review,” *The Journal of Supercomputing*, vol. 75, no. 5, pp. 2341–2394, 2019.
- [20] K. G. Omeke, M. S. Mollel, M. Ozturk et al., “DEKCS: a dynamic clustering protocol to prolong underwater sensor networks,” *IEEE Sensors Journal*, vol. 21, no. 7, pp. 9457–9464, 2021.
- [21] A. Farhat, C. Guyeux, A. Makhoul, A. Jaber, and R. Tawil, “On the coverage effects in wireless sensor networks based prognostic and health management,” *International Journal of Sensor Networks*, vol. 28, no. 2, pp. 125–138, 2018.
- [22] T. Wang, T. Liang, J. Fan, and X. Wei, “Sequential opening multi-jammers localisation in multi-hop wireless network,” *IET Information Security*, vol. 12, no. 5, pp. 445–454, 2018.