



## OPEN Wireless sensor network positioning technology based on improved sampling box and fuzzy reasoning

Ningyun Dan<sup>1</sup>, Lun Zhao<sup>1</sup>✉, Wenjin Pan<sup>2</sup> & Yan Cai<sup>1</sup>

As the wireless sensor network technology develops, precise positioning has become the key to achieving effective monitoring and data collection. This study proposes a wireless sensor network positioning technology based on improved sampling boxes and fuzzy reasoning to address the problem of low positioning accuracy and efficiency of traditional positioning methods in resource constrained environments. By using a fuzzy clustering algorithm based on time series optimization and a multidimensional Gaussian model to optimize the sampling box, the accuracy and efficiency of localization are significantly improved. The research results indicate that on the UCI and MIT Reality datasets, the Rand coefficient of the fuzzy clustering algorithm optimized for time series is stable at 0.89 when the weight allocation index is 5, which is superior to other comparative algorithms. When the proportion of beacon nodes reaches 40%, the average positioning error of the new model is as low as 0.21 m, which is lower than the comparison model. When the number of nodes is 50 and 200, the response times of the new model are 0.41 ms and 0.72 ms, respectively, which are also the fastest among all models. From this, the new model can maintain high positioning accuracy and fast response under different communication radii and beacon node ratios, especially in scenarios with a large number of nodes, demonstrating excellent versatility and efficiency. This study is of great significance for improving the application efficiency of wireless sensor networks in fields such as environmental monitoring and industrial automation, and provides an effective technical solution for achieving more accurate positioning.

**Keywords** Wireless sensor networks, Fuzzy reasoning, Multidimensional Gaussian model, Sampling box, Node positioning, Positioning accuracy, Communication radius

As the modern information technology develops, Wireless Sensor Networks (WSNs) have become one of the key technologies for achieving intelligent monitoring and automated control<sup>1</sup>. However, despite their broad applications in environmental monitoring and industrial automation, existing node localization techniques face critical challenges in resource-constrained environments (e.g., limited energy, computational power) and dynamic scenarios (e.g., multipath interference, non-line-of-sight propagation)<sup>2-4</sup>. The accuracy of node location directly affects the value of monitoring data and the performance of the network<sup>5</sup>.

Despite the remarkable advancements in WSN optimization methods concerning positioning accuracy, energy efficiency, and security, several limitations persist. The Kalman filter-based iterative boundary box algorithm, while effective in reducing positioning errors, suffers from heightened communication overhead due to its multi-loop optimization process. The unpredictable number of iterations also poses challenges for real-time applications. The small target detection algorithm, which enhances accuracy through optimized sampling boxes, is dependent on high-resolution image processing. This dependency, in turn, escalates node energy consumption. Although the gray wolf optimizer demonstrates improvements in intrusion detection rates, its multi-wolf coordination mechanism introduces additional calculation delays. The PSO positioning framework, while efficient in node displacement scenarios, necessitates a beacon node density of  $\geq 40\%$ , thereby increasing deployment costs significantly. The underwater multi-hop routing protocol, which prolongs network life via cluster head selection, fails to address the issue of multi-hop ranging error accumulation. Consequently, positioning errors increase exponentially with the number of hops. Ant colony optimization-

<sup>1</sup>College of Physics and Information Engineering, Zhaotong University, Zhaotong 657000, China. <sup>2</sup>Yunnan Electric Test & Research Institute Group Co., Ltd., Kunming 650000, China. ✉email: BEIbeizhu1232@163.com

based path planning, despite balancing energy consumption, experiences a substantial decrease in iterative convergence speed as the number of nodes increases. BIM modeling, although effective in improving signal prediction accuracy, demands that nodes possess 3D environment modeling capabilities, leading to increased hardware costs. Lastly, while HBGWO performs effectively in intelligent building coverage optimization, its dynamic parameter adjustment mechanism consumes considerable node memory resources. Collectively, these limitations result in diminished positioning accuracy and delayed responses, particularly in scenarios with constrained resources. To overcome these limitations and enhance the applicability and efficiency of WSNs in diverse, especially resource-constrained environments, a WSN positioning technology based on improved sampling boxes and fuzzy reasoning is proposed.

The innovation of the research mainly lies in the following aspects: (1) Combining Fuzzy Clustering Means based on Time series Pathology (FCMT) with multi dimensional Gaussian sampling boxes to improve positioning accuracy and efficiency. (2) Using fuzzy logic to handle uncertainty. (3) By improving the sampling box to enhance the ability to capture signal strength features and improve positioning accuracy. (4) Introducing a dynamic weight allocation mechanism in fuzzy reasoning to improve the adaptability and stability of the model. (5) Adopting nonlinear spatial modeling to reduce communication redundancy.

The main contributions of this article are as follows: (1) Proposing Fuzzy Clustering Means based on Time-series Morphology-Gauss Sampling box-WSN (FCMT-GS-WSN), Being able to accurately estimate node positions under resource constraints. (2) Under resource constraints, FCMT-GS-WSN can maintain high positioning accuracy while significantly reducing the model's response time and improving real-time performance. (3) Research on accurately simulating and predicting signal strength and node positions through 3D models. (4) Compared with the multi round iterative Kalman filter optimization method, the proposed fuzzy clustering and multi dimensional Gaussian sampling collaborative mechanism significantly reduces communication redundancy through nonlinear spatial modeling. (5) Compared with the dependency of the swarm intelligence optimization positioning framework on high-density beacon nodes, the study utilizes dynamic weight allocation through fuzzy reasoning to establish a more stable spatial mapping relationship under low beacon coverage, reducing deployment costs.

## Related work

In terms of improving the sampling box, Mani R et al. raised an iterative boundary box algorithm based on Kalman filtering to refine the estimated positions of unknown nodes and optimize the sampling box to address the node localization problem in large-scale WSNs. The experiment findings denoted that the new algorithm effectively reduced positioning errors (PEs) without requiring more equipment and increasing communication costs<sup>6</sup>. Lou H et al. proposed a small-sized object detection algorithm suitable for special scenes to optimize the sampling box and improve the detection accuracy of small-sized objects, to solve the problem of fatigue and judgment errors of traditional camera sensors when observing objects of different sizes for a long time in complex scenes. The experiment findings denoted that the detection accuracy of the new algorithm on the Visdrone dataset was superior to existing single view detection methods<sup>7</sup>. In terms of security optimization, Liu G et al. proposed an intelligent intrusion detection model for intrusion detection in WSNs and used Lévy flight strategy to adjust and optimize the sampling box. The experiment findings denoted that the proposed new algorithm performed well in benchmark function testing, and the intrusion detection model achieved 99% accuracy in performing Disk Operating System (DoS) intrusion detection. The proposed intrusion detection model had good performance<sup>8</sup>. Safaldin M et al. proposed an improved binary grey wolf optimizer support vector machine enhanced intrusion detection system to improve the intrusion detection capability of WSNs. The system optimized performance by adjusting the number of wolf packs, aiming to improve detection accuracy and rate, while reducing processing time and false alarm rate. The results showed that the new algorithm using seven wolves outperformed other algorithms in terms of accuracy, feature count, execution time, false positive rate, and detection rate<sup>9</sup>. Cao C et al. designed an improved identity encryption algorithm to solve the issues of limited energy, weak computing power, poor communication capability, and vulnerability in WSNs. The findings indicated that the new algorithm simplified the key generation process, reduced network traffic, and improved network security. It had low energy consumption and high security, and was suitable for WSN applications with high security requirements<sup>10</sup>.

In terms of WSN localization and energy consumption optimization, Kanwar V et al. designed a Distance Vector-Hop localization framework based on particle swarm optimization to solve the problem of node displacement due to disasters and other events in WSNs. The experimental results showed that compared with traditional positioning frameworks, the proposed positioning framework performed better in reducing time consumption, PEs, and energy consumption<sup>11</sup>. Kathirolu et al. proposed a hybrid sparrow search algorithm based on differential evolution algorithm to solve the problems of limited battery power and high energy consumption in WSNs. This method extended the life cycle of the sampling box by selecting cluster heads and using an energy allocation mechanism. The outcomes indicated that the algorithm had significant improvements in residual power and throughput, effectively improving the lifespan of nodes<sup>12</sup>. Lakshmana K et al. proposed an improved meta heuristic driven energy-saving clustering routing scheme to address energy efficiency issues in IoT assisted WSNs. This scheme used an improved Archimedes optimization algorithm for cluster head election and organization. The outcomes denoted that the new method outperformed existing methods in terms of network lifecycle, energy consumption, packet transmission rate, and latency<sup>13</sup>. Mohan P proposed an improved clustering technique based on meta heuristic multi-hop routing protocol to address energy efficiency issues in underwater WSNs. The outcomes indicated that the new technology could significantly improve the energy efficiency and network lifetime of underwater WSNs, and was superior to existing methods in multiple performance indicators<sup>14</sup>. Roy S et al. proposed a mobile aggregation node path optimization method based on ant colony optimization algorithm to solve the problems of uneven energy consumption and data transmission

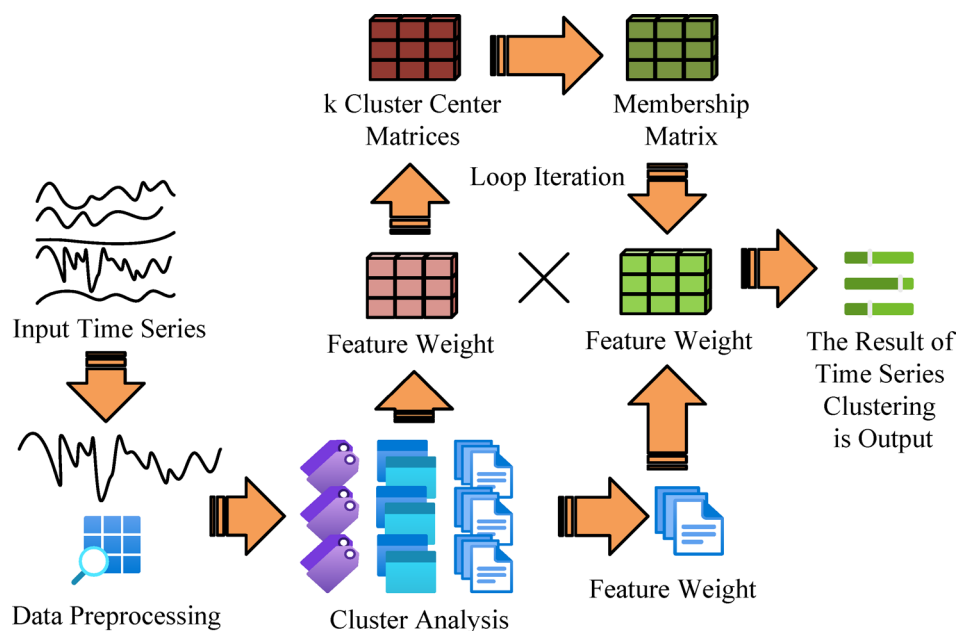
delay in WSNs. The simulation results showed that this protocol outperformed existing methods in terms of network lifespan, energy consumption, packet transmission rate, and end-to-end latency<sup>15</sup>. A Ahmadi et al. artificially accurately predict the network channel characteristics under the occlusion of buildings, and establish an indoor wireless network based on Building Information Modeling (BIM), electromagnetic simulation and measurement. Experimental results show that the new method can extract network channel features accurately<sup>16</sup>. A Chambon et al. proposed Pareto-Envelope Selection based on Non-Dominated Sorting Genetic Algorithm (PREP-NSGA) to optimize indoor deployment of IoT for efficient integration of IoT devices. The results show that the coverage of the proposed scheme under connection constraints is better than that of existing schemes and random deployment<sup>17</sup>. K Zaimen et al. proposed Hybrid Binary Grey Wolf Optimiser (HBGWO) to reduce the cost of wireless sensor network coverage for smart buildings. The experimental results show that HBGWO is superior to existing methods in terms of network coverage and deployment cost under connectivity constraints<sup>18</sup>. R P SHada et al. proposed a hybrid positioning technology integrating random forest and multilateral measurement to solve the problem of difficult GPS positioning caused by WSN. Experiments show that this method is effective in both simulation and real prototypes<sup>19</sup>. T T Nguyen et al. proposed innovative Adaptive Optimisation Approach Called Crossover Mutation Differential Evolution (ACMDE), to address the high computational complexity of traditional node localization methods in large-scale WSN. The experimental results show that ACMDE performs better than other algorithms in node localization tasks<sup>20</sup>.

## Method

### Improvement of fuzzy reasoning algorithm

In WSNs, positioning technology is crucial for achieving precise monitoring and data collection<sup>21</sup>. However, due to resource limitations and environmental factors of sensor nodes, traditional positioning methods often face challenges in accuracy and efficiency<sup>22</sup>. Among them, fuzzy reasoning, as an effective method for dealing with uncertainty and nonlinear problems, originated from fuzzy set theory and has been widely used to improve the positioning accuracy of WSN<sup>23</sup>. Fuzzy sets allow elements to have different degrees of membership within the set, which enables fuzzy reasoning to better handle ambiguity and uncertainty in the real world<sup>24</sup>. Fuzzy reasoning systems form an effective information processing mechanism by fuzzifying inputs, applying fuzzy rules, and defuzzifying outputs<sup>25</sup>. Based on this, the study improved the fuzzy clustering algorithm and proposed a Fuzzy Clustering Means based on Time-series Morphology (FCMT) based on time series optimization. The structure of FCMT algorithm is shown in Fig. 1.

In Fig. 1, the FCMT algorithm is a clustering method specifically designed for time series data. In order for each time series data to have an initial classification, the FCMT algorithm first initializes the number of clusters and randomly assigns initial membership degrees to each time series data. Then, to capture the similarities between time series, Jeffrey's composite distance is used to measure the distance between time series. This distance measurement method has linear time complexity and can effectively handle real valued time series data. During the iteration process, to improve the quality of clustering, the algorithm adaptively learns the feature weights of each cluster's time series based on its membership degree, to capture the unique core features of each cluster and weight the distances accordingly. Therefore, the time series similarity forms corresponding to the core features of each cluster become the core elements in the cluster partitioning process. As the iteration progresses, to more accurately reflect the underlying structure of the data, the algorithm continuously updates the cluster center and membership matrix until the stopping criterion is met, that is, the change in the objective



**Fig. 1.** FCMT algorithm structure.

function is less than the predetermined threshold or reaches the maximum number of iterations. Finally, to determine which cluster each time series data ultimately belongs to, the time series is clustered according to the principle of maximum membership degree, and the clustering results are obtained. The process of Jeffrey's composite distance measurement is shown in Fig. 2.

In Fig. 2, Jeffrey's composite distance is a similarity measurement method designed for time series data, which measures the distance between them by considering the morphological features of the time series. The process first preprocesses the time series data to ensure that it is a positive real number, and then applies the adaptive translation Jeffrey's distance formula to measure the differences between time series feature points. The formula for Jeffrey's composite distance is shown in Eq. (1)<sup>26</sup>.

$$J_{\text{comp}}(y_i^p, c_l^p) = (y_i^p - c_l^p) [\ln (y_i^p + [|D_{\text{min}}|]) - \ln (c_l^p + [|D_{\text{min}}|])] \tag{1}$$

In Eq. (1),  $y_i^p$  means the value of the  $i$  th time series on the  $p$  th feature,  $c_l^p$  means the value of the center point of cluster  $l$  on the  $p$  th feature,  $D_{\text{min}}$  represents the minimum value of the entire time series dataset, and  $[|D_{\text{min}}|]$  represents the absolute value of  $D_{\text{min}}$  rounded up. Jeffrey's composite distance can help FCMT algorithm more accurately capture the similarities and differences of time series. The calculation method of autocorrelation coefficient for time series is shown in Eq. (2)<sup>27</sup>.

$$\rho(b) = \frac{\sum_{i=1}^{n-b} (y_i - \bar{y})(y_{i+b} - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{2}$$

In Eq. (2),  $\rho(b)$  is the autocorrelation coefficient of the lagged  $b$  step of the time series,  $y_i$  is the value of the time series at time  $i$ , and  $\bar{y}$  is the mean of the time series. First order moving averages usually do not involve lag steps, and in higher-order cases, the calculation of time series is shown in Eq. (3)<sup>28</sup>.

$$y_i = \Lambda + \epsilon_i + \theta_1 \epsilon_{i-1} + \theta_2 \epsilon_{i-2} \tag{3}$$

In Eq. (3),  $\Lambda$  is the mean,  $\epsilon_i$  is the white noise,  $\theta_1$  and  $\theta_2$  are the white noise coefficients. The formula for simple exponential smoothing is shown in Eq. (4).

$$\hat{y}_{i+1|i} = \alpha y_i + (1 - \alpha) \hat{y}_{i|i-1} \tag{4}$$

In Eq. (4),  $\hat{y}_{i+1|i}$  is the prediction of  $y_{i+1}$ ,  $\alpha$  is the smoothing parameter,  $i + 1$  represents the next moment, and  $i - 1$  represents the previous moment. The distance calculation method between time series and cluster centers is shown in Eq. (5)<sup>29</sup>.

$$d_{i,l} = \sum_{p=1}^z J_{\text{comp}}(y_i^p, c_l^p) w_{l,p} \tag{5}$$

In Eq. (5),  $d_{i,l}$  means the distance between the time series  $i$  and the cluster center  $l$ , and the weight of the  $p$  th feature in the  $w_{l,p}$  point cluster  $l$ . The calculation method of feature weights is shown in Eq. (6)<sup>30</sup>.

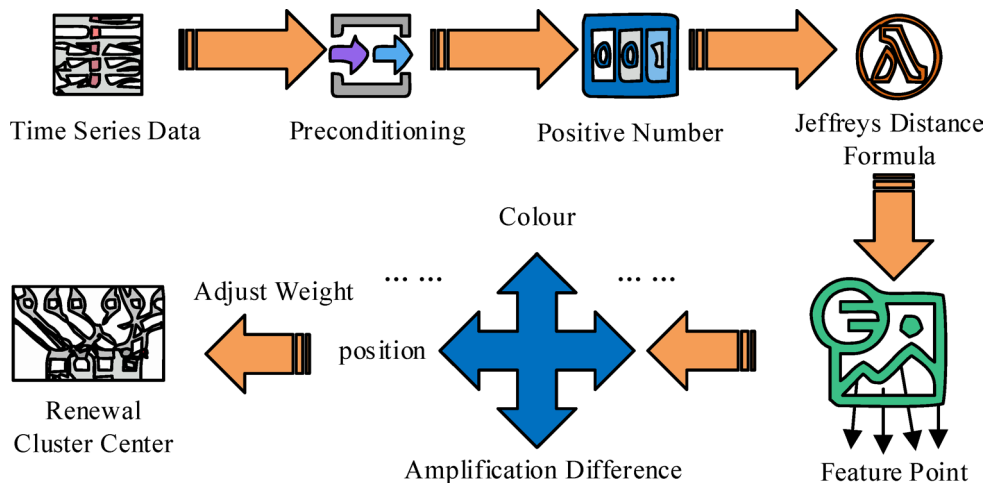


Fig. 2. Jeffrey's composite distance measurement process.

$$w_{l,p} = \left( \frac{\sum_{i=1}^n \mu_{i,l}^\beta J_{\text{comp}}(y_i^p, c_l^p)}{\sum_{p=1}^z \sum_{i=1}^n \mu_{i,l}^\beta J_{\text{comp}}(y_i^p, c_l^p)} \right)^{\frac{1}{\alpha-1}} \quad (6)$$

In Eq. (6),  $\mu_{i,l}$  represents the membership degree of time series  $i$  to cluster  $l$ , and  $\beta$  represents the fuzzy coefficient used to control the degree of fuzziness of the membership degree.  $\alpha$  stands for weight allocation index, used to adjust the distribution of feature weights. The formula for calculating membership degree is shown in Eq. (7)<sup>31</sup>.

$$\mu_{i,l} = \frac{1}{\sum_{k=1}^c \left( \frac{d_{i,l}}{d_{i,k}} \right)^{\frac{2}{\beta-1}}} \quad (7)$$

In Eq. (7),  $k$  represents the cluster index during the clustering process. The calculation method of the objective function is denoted in Eq. (8).

$$J(U, V) = \sum_{i=1}^n \sum_{l=1}^c \mu_{i,l}^\beta d_{i,l}^2 \quad (8)$$

In Eq. (8),  $J(U, V)$  represents the objective function applied to assess the quality of clustering. The formula for updating cluster centers is shown in Eq. (9).

$$c_l = \frac{\sum_{i=1}^n \mu_{i,l}^\beta y_i}{\sum_{i=1}^n \mu_{i,l}^\beta} \quad (9)$$

In Eq. (9),  $c_l$  represents the updated cluster center. The formula takes into account the differences between each feature point in the time series and its corresponding cluster center feature point, and amplifies these differences through a natural logarithm function to more sensitively capture the morphological changes in the time series. Applying Brownian motion model to FCMT algorithm can make the distance measurement more consistent with the actual wireless signal propagation law. This method modifies the distance measurement standard of FCMT algorithm by simulating the change of signal strength with time, so as to improve the indoor positioning accuracy. During the iteration process, the algorithm utilizes membership information to dynamically adjust feature weights, which reflect the importance of different features in the clustering process. Finally, the weighted distance values are used to update the membership degrees of the time series to each cluster, thereby updating the cluster centers. This process is iteratively repeated until the stopping criterion is met.

### Improvement of sampling box

After introducing the improvement of fuzzy reasoning algorithm, the study will further explore how to raise the positioning accuracy of WSNs by optimizing the sampling box. The sampling box is a basic unit used to collect information in wireless sensor networks. The collected information can be uploaded to the server through wireless transmission. The FCMT algorithm can effectively process time series data and improve the positioning accuracy in WSNs. This algorithm captures the morphological features of time series by adaptively learning their feature weights, thereby optimizing the clustering process. In WSNs, the sampling box serves as the basic unit for collecting information, and its performance directly affects the monitoring and positioning capabilities of the entire network. To further raise the accuracy and efficiency of positioning, a multidimensional Gaussian model is studied to optimize the sampling box. The multidimensional Gaussian model can model the relationship between the signal strength received by the sampling box and the node position, providing a probability density function for each node. Multidimensional Gaussian models help network systems to more accurately estimate the positions of nodes, improve the sampling quality of sampling box nodes, and thereby enhance the positioning capability of the entire network. Fuzzy reasoning can deal with uncertainty and nonlinear problems in the localization process, while Gaussian modeling accurately describes the relationship between signal strength and node position through probability density function. The integration will effectively improve the accuracy and robustness of positioning. The structure of the multidimensional Gaussian model is indicated in Fig. 3.

In Fig. 3, the multidimensional Gaussian model, also known as the multivariate normal distribution, is a commonly used model in probability theory and statistics to describe the joint probability distribution of multiple random variables. This model assumes a linear relationship between variables and a symmetrical distribution. The model assumes that the relationship between the signal strength and the node position can be described by a probability distribution, that the propagation follows a lognormal distribution, and that the signal strength at the path loss index and reference distance is known. The core of the model is the mean vector and covariance matrix. The mean vector denotes the expected value of a random variable, pointing to the center position of the dataset in multidimensional space, while the covariance matrix describes the correlation and variability between variables, with diagonal elements representing the variance of variables and non diagonal elements representing the covariance between variables. The probability density function of a multidimensional Gaussian distribution is given by a complex exponential function, which includes the inverse of the mean vector

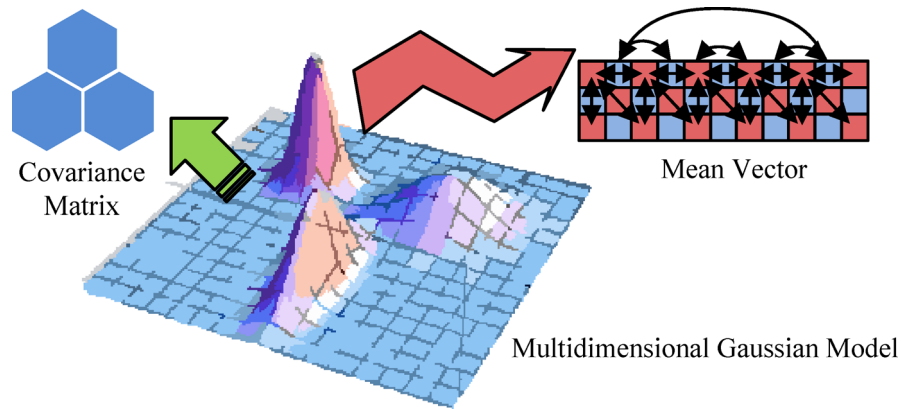


Fig. 3. Multidimensional Gaussian model structure.

and covariance matrix, as well as the quadratic form of the difference between the variable vector and the mean vector. This function can quantify the probability density of a given set of observations. In WSN positioning, multidimensional Gaussian models can be used to simulate the relationship between node Received Signal Strength (RSS) and location. By optimizing model parameters, the accuracy and reliability of positioning can be raised<sup>32</sup>. This article studies the application of Building Information Modeling (BIM) to construct an accurate three-dimensional model of the building structure, so as to simulate and predict the relationship between signal strength and node position more accurately, provide a more accurate probability density function, and thus improve the sampling quality of sampling box nodes. The calculation method for signal strength path loss from the sampling box is shown in Eq. (10)<sup>33</sup>.

$$P_M(f) = P_M(f_0) + 10m \log_{10}(f/f_0) + Y_\omega \tag{10}$$

In Eq. (7),  $P_M(f)$  represents the path loss at the distance  $f$ ,  $P_M(f_0)$  represents the path loss at the reference distance  $f_0$ ,  $m$  is the path loss index, and  $Y_\omega$  means a random variable representing noise. The calculation method for the average RSS vector of the reference point is shown in Eq. (11)<sup>34</sup>.

$$P_{rss} = \left( \frac{1}{N} \sum_{j=1}^N P_{rss,j} \right) \tag{11}$$

In Eq. (11),  $P_{rss}$  represents the average RSS vector of the reference point, and  $P_{rss,j}$  means the signal strength of the  $j$ th anchor node (AN) received by the reference point. The calculation method for RSS mapping maps is shown in Eq. (12)<sup>35</sup>.

$$B = (b_{kz})_{p \times q} \tag{12}$$

In Eq. (12),  $B$  represents the RSS mapping map, and  $b_{kz}$  represents the signal strength received by the  $z$ th AN at the  $k$ th reference point. By measuring RSS, it is possible to determine which nodes are within each other's communication range, thereby helping to construct the topology of the network. The formula for the multidimensional Gaussian distribution of ideal and reference points is shown in Eq. (13).

$$G(y; v, \Omega) = \frac{1}{(2\pi)^{e/2} |\Omega|^{1/2}} \exp \left( -\frac{1}{2} (y - v)^T \Omega^{-1} (y - v) \right) \tag{13}$$

In Eq. (13),  $G(y; v, \Omega)$  represents a multidimensional Gaussian distribution function used to describe the similarity between the ideal point  $y$  and the reference point  $v$ , and  $T$  represents transpose. The correction formula for the influence of the adjustment function on the positional relationship is shown in Eq. (14).

$$b^* = \arg \min_b \text{MSE} \tag{14}$$

In Eq. (14),  $b^*$  represents the adjustment function factor. Finding the optimal adjustment function factor minimizes the mean square error MSE, thereby correcting the distance relationship deviation between the ideal point and the node to be located. The node position estimation formula is shown in Eq. (15).

$$\tilde{Q} = \frac{\sum_{k=1}^M U_k \cdot V_k}{\sum_{k=1}^M U_k} \quad (15)$$

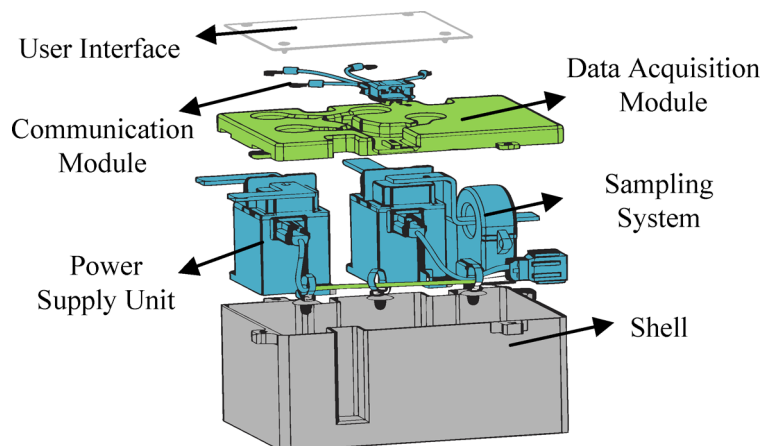
In Eq. (15),  $\tilde{Q}$  represents node position,  $U_k$  represents position similarity,  $V_k$  represents reference point position, and  $M$  means the total amount of reference points involved in position estimation. Based on this, the improved sampling box structure is shown in Fig. 4.

In Fig. 4, the sampling box has a sturdy casing to protect the internal sensor unit. It collects data through a data collection module and transmits the information to a remote server via wireless network. The device is powered by a power supply unit and equipped with a user interface for operation and status display. In the sampling box, a multidimensional Gaussian model is used for anomaly detection, data fusion, and feature extraction to identify outliers in the data and improve data quality. By using Gaussian process for spatiotemporal multi-sensor calibration, more accurate target position determination can be achieved. The sampling frequency of the sensor node is not less than 100 Hz, the expansion range is 0.3–1.5 m, the accuracy is 0.5%, and the Hall sensor redundancy sampling is used to control the comprehensive error below  $\pm 0.58\%$ . The improvements of the sampling box is reflected in the intelligent calibration and modular structure design. The multi-dimensional Gaussian model of the sampling box can realize multi-sensor space-time calibration and anomaly detection, improve data accuracy, and the split modular architecture can reduce maintenance costs. The architecture of a sample box usually consists of four parts. The physical layer contains protective housings, sealed doors, and internal partitions that provide mechanical support and environmental isolation. The sampling module integrates sensor and sampling mechanism to realize accurate sample acquisition. The storage control layer includes an electric drive, a valve control system, and a sample staging chamber to ensure the automation of the sampling process and sample stability. At the same time, the data interaction layer is configured with barcode identification, communication interface and data transmission module, which supports sample information tracking and remote monitoring.

#### Establishment of Wsn localization model

On the basis of successfully improving the fuzzy clustering algorithm and sampling box, a comprehensive WSN positioning model will be constructed. In WSNs, the FCMT algorithm establishes a fuzzy mapping between signal reception strength and distance, enabling the system to more accurately estimate node positions in noisy environments. The sampling box creates a more accurate sampling area by combining the position information of ANs and ordinary nodes, thereby reducing the search space, minimizing invalid samples, and improving positioning accuracy and sampling efficiency. The improvement of fuzzy reasoning algorithm and sampling box plays an important role in enhancing the positioning accuracy of WSNs. Based on this, a new WSN model has been developed. The structure of the WSN is shown in Fig. 5.

In Fig. 5, WSNs are self-organizing networks composed of a large number of sensor nodes, which are connected through wireless communication and work together to monitor the environment or perform tasks. The structure of WSNs typically includes ordinary sensor nodes, aggregation nodes, and management nodes. Ordinary sensor nodes are the main components of a network, responsible for collecting environmental data such as temperature, humidity, light intensity, etc., and forwarding the data to aggregation nodes<sup>36</sup>. These nodes typically have limited resources, including computing power, storage space, and communication range. The aggregation node is an enhanced sensor node with stronger computing, storage, and communication capabilities<sup>37</sup>. They act as gateways, responsible for relaying data from sensor nodes to management nodes, while also conveying instructions from management nodes to regular sensor nodes. A management node is typically a central computer or console responsible for displaying network collected data, publishing monitoring tasks, and



**Fig. 4.** Improved sampling box structure.

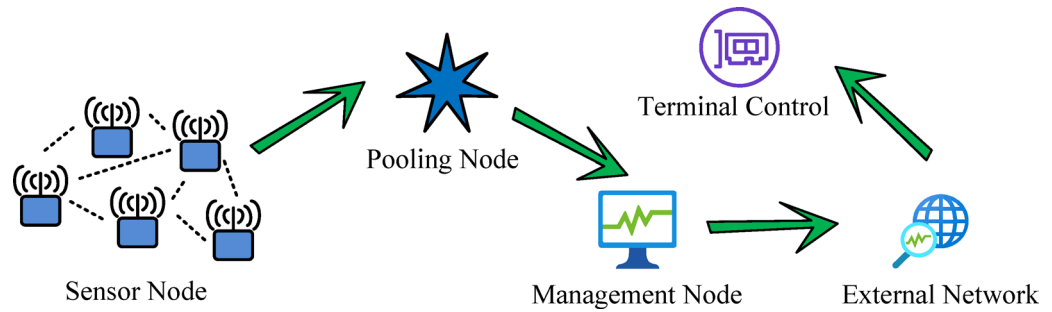


Fig. 5. WSN architecture.

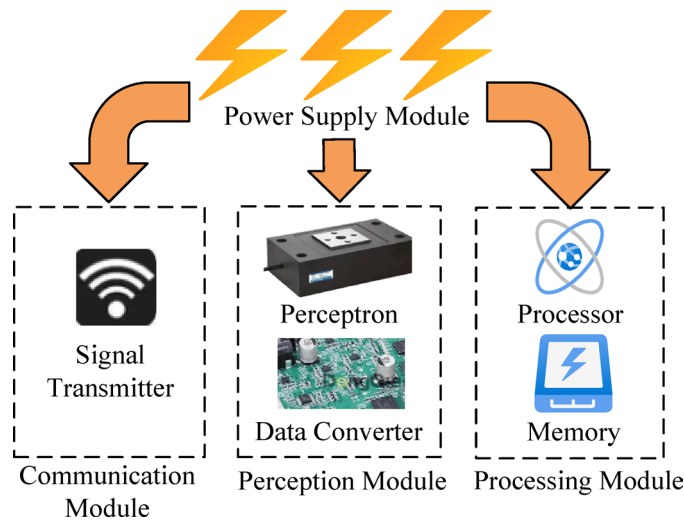


Fig. 6. Sensor node structure.

performing data processing. It can connect to external networks to achieve further analysis and storage of data. The nodes of WSNs are usually randomly deployed in the monitoring area, forming a multi hop communication network<sup>38</sup>. The communication between nodes depends on their communication radius, and there is usually direct communication or relay between nodes through other nodes. Although the multi-hop communication improves the coverage and flexibility of the network, the topology of the network may change dynamically due to the mobility of nodes, energy depletion or environmental changes, which challenges the stability and reliability of the network. To address this challenge, this article presents a wireless sensor network localization technology based on improved sample box and fuzzy reasoning, which can adapt to the changes of network topology, especially in the evaluation of data sets including mobile devices. The key technologies of WSNs include topology control, node localization, network protocols, time synchronization, and data fusion. The Ultra-Wideband (UWB) positioning system has the characteristics of high precision and low power consumption, which is suitable for scenarios requiring high precision positioning. Therefore, the UWB positioning system is used for data correlation to reduce the multipath effect of wireless sensor network positioning model and the positioning accuracy in non-line-of-sight environment. These technologies collectively support the efficient and stable operation of WSNs in various application scenarios. The structure of the sensor node is denoted in Fig. 6.

In Fig. 6, sensor nodes are the basic units of WSNs, consisting of several key modules: power supply module, communication module, perception module, and processing module. The power supply module provides energy to nodes, usually powered by batteries. Considering that nodes may be deployed in unmanned environments for a long time, the power supply module often needs to have low power consumption characteristics. The communication module is responsible for wireless data transmission between nodes, enabling nodes to communicate with each other or with base stations. The perception module is the interface for nodes to interact with the external environment and collect environmental data. These sensors convert physical phenomena into electrical signals, providing monitoring data for the network. The processing module is the brain of the node, which typically includes a microprocessor and memory for performing data preprocessing, decision making, and basic control tasks. It may also be responsible for running positioning algorithms, processing data from perception modules, and coordinating with other nodes or base stations. In addition, the security module in sensor nodes is used to protect the integrity and privacy of data. The external interface module of the sensor node is used for data exchange with other devices or networks. The design of nodes emphasizes miniaturization,

low cost, and high energy efficiency to satisfy the demands of large-scale deployment. The entire process of WSN positioning technology is shown in Fig. 7.

In Fig. 7, the process of WSN positioning technology first requires network deployment, where sensor nodes are dispersed and arranged in the target monitoring area. These nodes include ANs and ordinary nodes to be located. ANs know their own location, while ordinary nodes need to determine their location through positioning algorithms. During the initialization phase, the network forms a self-organizing structure and ANs begin broadcasting their location information. In the data collection stage, ordinary nodes collect environmental data and information such as signal strength from ANs. Using this information, the data is processed using a multidimensional Gaussian distribution. Next, data fusion technology is utilized to integrate multi-source information and optimize location estimation. Then the FCMT algorithm is executed to calculate the location information. Finally, the calculated location information is confirmed and applied. The pseudocode of the FCMT-GS-WSN model is shown in Fig. 8.

## Result

### Analysis of clustering effect of Fcmt algorithm

To ensure the positioning accuracy of WSNs, the study adopted the UCI Machine Learning Repository and MIT Reality dataset. The UCI dataset provides a rich machine learning dataset containing multiple types of data features that can be used to model various environmental parameters in the WSN, such as temperature, humidity, etc. The MIT Reality dataset is a large-scale dataset used for researching and developing indoor positioning systems, which contains a large number of wireless signals. The ratio of training set to testing set is 80%: 20%. The structure, accuracy, and integrity of the dataset are directly determined by the hardware architecture of the sampling box and the deployment strategy of the sensor nodes. The modular design of the sampling box determines the data generation hierarchy, while the spatiotemporal distribution of nodes affects the data coverage range. The experimental platform parameters are indicated in Table 1.

According to Table 1, the study constructed an experimental platform consisting of 100 nodes, with 20% serving as ANs, randomly deployed in an indoor area of 500 square meters. The data collection frequency was set to once per second, and each experiment lasted for 24 h. In terms of algorithm parameters, 5 clusters were preset, initial membership degrees were randomly assigned, the fuzzy coefficient  $m$  was set to 2, and the weight allocation index  $q$  was also set to 2. The stopping criterion for the algorithm was that the change in the objective function was less than 0.001 or the number of iterations reached 1000. In addition, the study used Jeffrey's composite distance to measure time series data and optimized the sampling box through a multidimensional Gaussian model, with a path loss index set to 2, the radial basis function (RBF) kernel length scale is set to 0.5. Improved quartile range (IQR) based method for anomaly detection in the sampling box with an IQR coefficient of 1.5 to determine the threshold for outliers. The comparative algorithms for analyzing the clustering performance of FCMT algorithm were R-Clustering, Two-stage Statistical Segmentation-clustering Time Series Procedure (TS3C), K-Means clustering algorithm, and Biclustering using Optimal Transport (BCOT). The

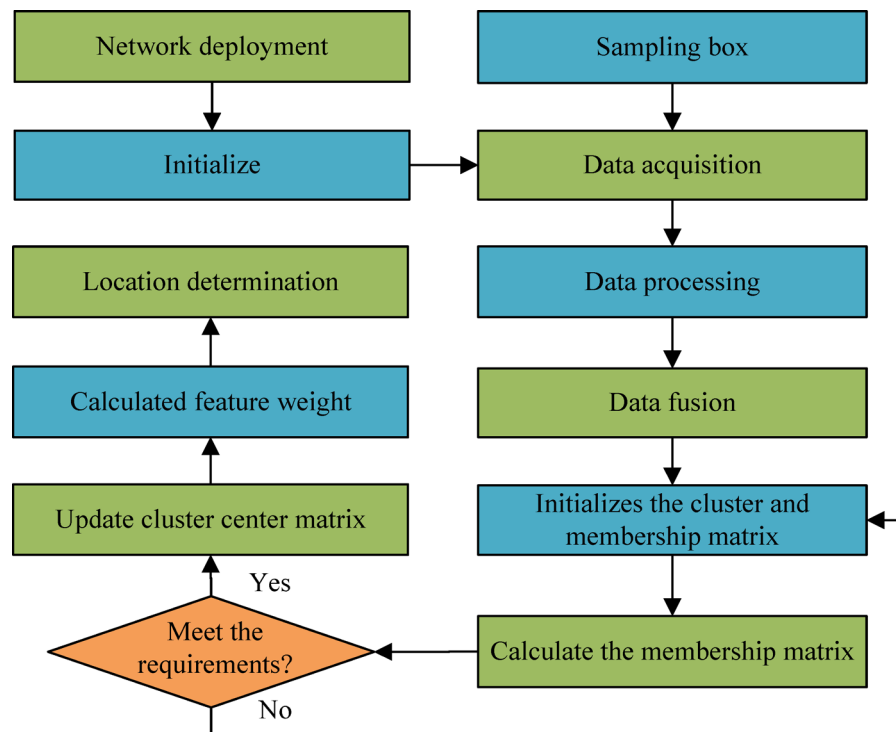


Fig. 7. WSN positioning technology flow chart.

```

// FCMT-GS-WSN Localization (Optimized)
Input: N=100, A=20%N, R=40m, C=5, m=2.0, q=5, ε=0.001
Output: Node positions

Begin:
1. Preprocess RSS time-series:
  • Gaussian filter → Normalize [0,1] for all unknown nodes

2. FCMT Clustering:
  • Initialize: U(random), V(kmeans_plus_plus)
  • While (iter++ < MAX_ITER && ΔJ > ε):
     $d_{\{i,k\}} = \sum [ (x_{\{i,f\}} + x_{\min}) \cdot \ln((x_{\{i,f\}} + x_{\min}) / (v_{\{k,f\}} + x_{\min})) ]$  // Eq.1 [1](@ref)
     $w_{\{k,f\}} = [\sum (u_{\{i,k\}}^m \cdot d^2)]^{-1}$  // Eq.6 [1](@ref)
     $u_{\{i,k\}} = 1 / \sum_j (d_{\{i,k\}} / d_{\{i,j\}})^{2/(m-1)}$  // Eq.7 [1](@ref)
     $v_{\{k,f\}} = \sum (u_{\{i,k\}}^m \cdot x_{\{i,f\}}) / \sum u_{\{i,k\}}^m$  // Eq.9 [1](@ref)
     $J = \sum \sum u_{\{i,k\}}^m \cdot d_{\{i,k\}}^2$  // Eq.8 [1](@ref)

3. Position Estimation:
  • Load BIM_3D_model // Sec 3.2 [7](@ref)
  • For each unknown node:
    Candidate_points = get_points_in_cluster(Cluster_labels[u])
     $P(r|RSS_u) \propto \exp(-0.5(RSS_u - \mu_r)^T \Sigma_r^{-1}(RSS_u - \mu_r))$  // Eq.13 [1](@ref)
    Optimized_position =  $\sum [S(r) \cdot \text{Coordinates}(r)]$  for  $r \in \text{Candidates}$  // Eq.15 [1](@ref)
    Final_position = Optimized_position +  $\delta \cdot (RSS_u - \mu_r)$  // Eq.14 [1](@ref)

Return Final_position, Positioning_Error, Response_Time
End

```

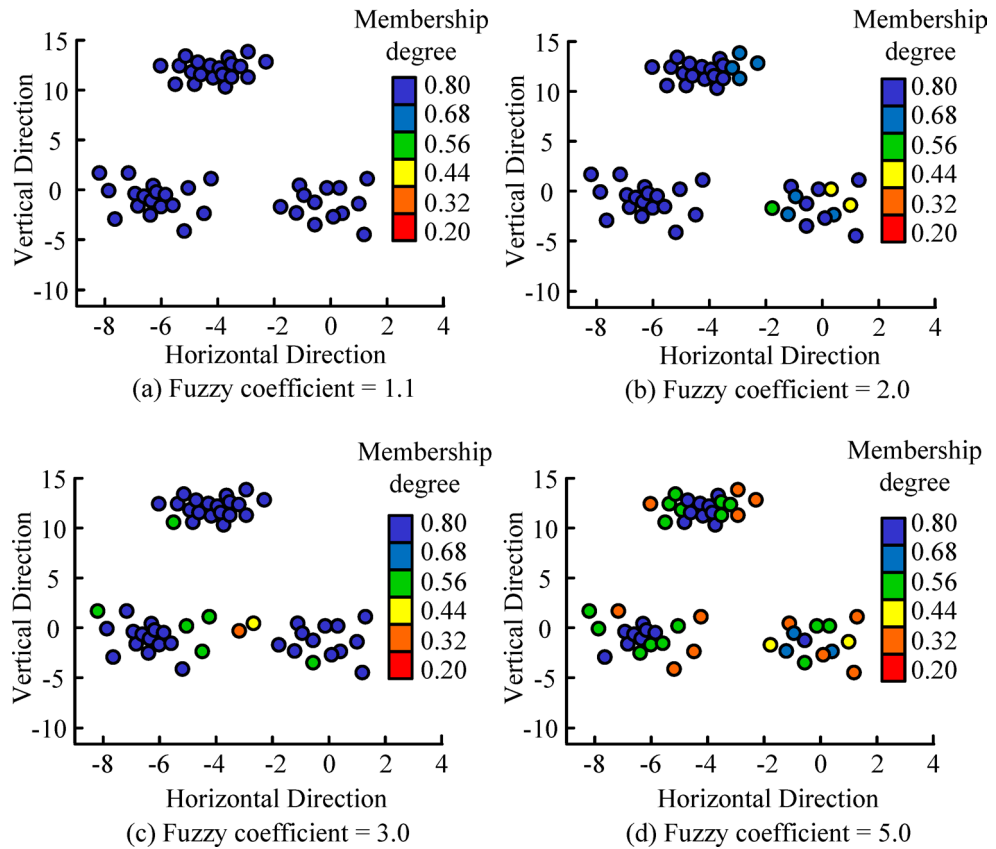
**Fig. 8.** The pseudocode of the FCMT-GS-WSN model.

Number of nodes	100
Interior area	500 m <sup>2</sup>
Communication radius	50 m
Node processor	ARM Cortex-M series
Node RAM	128 KB
Flash storage	1 MB
Expected working life	More than 3 years

**Table 1.** Experimental platform parameter table.

evaluation index was Rand Index (RI), which ranges from 0 to 1. The closer it is to 1, the better the clustering performance of the algorithm. Different fuzzy coefficients affect the performance of the algorithm, and the clustering outcomes of the FCMT algorithm under different fuzzy coefficients are shown in Fig. 9.

In Fig. 9 (a), when the fuzzy coefficient was 1.1, the clustering outcomes of FCMT algorithm tended to be hard clustering, and the membership degree of data points tended to be 1. The characteristics of fuzzy clustering were not fully reflected, resulting in overly clear clustering boundaries and a lack of the gradient characteristics that fuzzy clustering should have. In Fig. 9 (b), when the fuzzy coefficient was 2, the FCMT algorithm provided a more balanced clustering effect, and the membership distribution of data points was reasonable, which can better reflect the fuzzy clustering characteristics of the data, making the clustering results have both fuzzy clustering characteristics and good discriminability. In Fig. 9 (c), when the fuzzy coefficient increased to 3, although the fuzziness was enhanced, the membership degree of the boundary data points was too low, resulting in unclear clustering attribution. This reflected that the interpretability of clustering results decreased when the fuzzy coefficient was high. In Fig. 9 (d), when the fuzzy coefficient was further increased to 5, the membership degree of most data points was very low, and the clustering results were difficult to judge. This indicated that excessively high fuzzy coefficients would lead to overly fuzzy clustering effects, which was not conducive to accurately identifying the clustering attribution of data points. In summary, when the fuzzy coefficient was 2, the FCMT algorithm could best reflect the characteristics of fuzzy clustering while maintaining the clarity



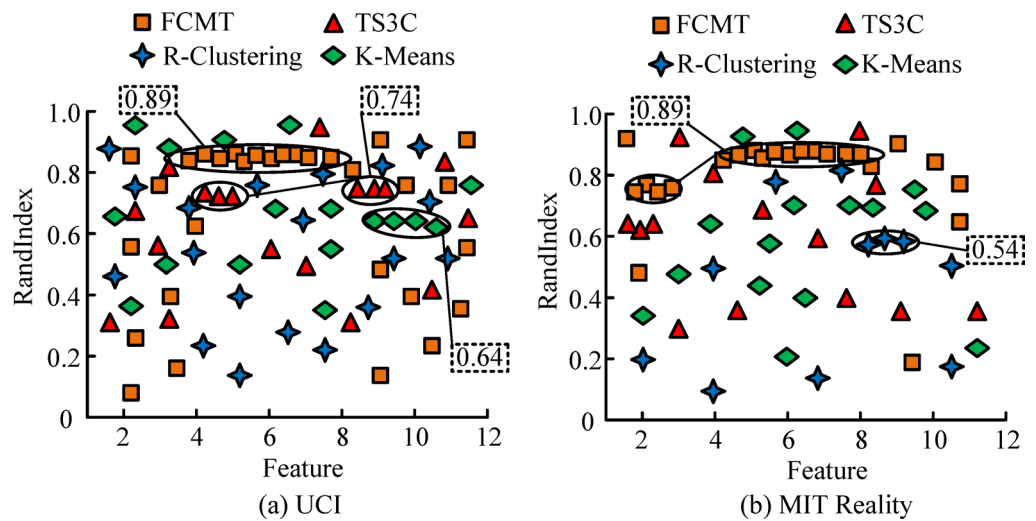
**Fig. 9.** Clustering results of FCMT algorithm with different fuzzy coefficients.

Weight distribution index	FCMT	TS3C	R-Clustering	K-Means	BCOT
1	0.21	0.18	0.19	0.17	0.18
2	0.65	0.51	0.39	0.38	0.42
3	0.78	0.62	0.48	0.51	0.58
4	0.86	0.69	0.51	0.60	0.63
5	0.89	0.74	0.52	0.64	0.61
10	0.89	0.74	0.53	0.64	0.61
15	0.89	0.74	0.54	0.64	0.61
20	0.89	0.74	0.54	0.64	0.61
25	0.89	0.74	0.54	0.64	0.61
30	0.89	0.74	0.54	0.64	0.61

**Table 2.** RI of the algorithm under different weight allocation indices.

and interpretability of clustering results, achieving the best balance between clustering effect and fuzziness. This moderate degree of fuzziness helps the FCMT algorithm to obtain more accurate clustering results while maintaining a certain degree of fuzziness in the data points. When the fuzzy coefficient was 2, the impact of the weight allocation index on the clustering performance of the algorithm is denoted in Table 2.

According to Table 2, the RI of the FCMT algorithm stabilized at 0.89 when the weight allocation index was 5. The RI of the TS3C algorithm stabilized at 0.74 when the weight allocation index was 5. The RI of the R-Clustering algorithm stabilized at 0.54 when the weight allocation index was 15. The RI of the K-Means algorithm stabilized at 0.64 when the weight allocation index was 5. From this, the RI of the algorithm increased with the increase of the weight allocation index, but when the weight allocation index increased to a certain range, the RI of the algorithm would also stabilize. Among them, the FCMT algorithm was the fastest among all algorithms to reach a stable state, and its RI was also the highest among all algorithms, demonstrating good clustering performance and robustness. When the fuzzy coefficient was 2 and the allocation index was 15, the clustering performance of each algorithm in the UCI and MIT Reality datasets is shown in Fig. 10.



**Fig. 10.** The clustering effect of each algorithm on different datasets.

Performance index	FCMT [This study]	GAECH <sup>39</sup>	WSA <sup>40</sup>
Average positioning error (m)	0.21	0.38	0.45
Response time (ms)	0.72	1.05	1.2
Average daily energy consumption per node (mJ)	8.75	12.3	14.6
Network life cycle (Rounds)	1520	980	850
Number of convergence iterations	26	42	38

**Table 3.** Performance comparison table of FCMT, GAECH and WSA.

In Fig. 10 (a), due to the optimization of the FCMT algorithm in terms of time series, it had a good clustering effect on features 4–8 on the UCI dataset, with an RI of 0.89. The TS3C algorithm exhibited clustering in features 5 and 9, with an RI value of 0.74. The clustering effect of R-Clustering algorithm was not significant. The K-Means algorithm exhibited clustering in features 9–10, but the RI value was relatively low at 0.64. In Fig. 10 (b), the FCMT algorithm still maintained good clustering performance on the MIT Reality dataset, with good clustering performance in features 2–3 and 5–8, and an RI of 0.89. The clustering performance of other algorithms was poor, although R-Clustering has shown clustering effect, the RI value was only 0.54. From this, it can be seen that the FCMT algorithm can play a stable clustering role in different environments and has strong universality. Genetic Algorithm for Energy-Efficient Clustering and Node Placement (GAECH) and Water Strider Algorithms (WSA) It has also been compared as a comparative algorithm<sup>39,40</sup>. Its performance is shown in Table 3.

According to Table 3, the positioning error of FCMT (0.21 m) is significantly lower than that of GAECH (0.38 m) and WSA (0.45 m), mainly due to the accurate modeling of signal propagation by the multidimensional Gaussian model. The response speed of FCMT is 0.72 ms, which is fast. The network has a lifecycle of 1520 rounds and a convergence iteration of 26 times, all of which are better than the comparison algorithms. It can be seen that the overall performance of FCMT is good and can meet the real-time requirements of high-precision positioning.

Fuzzy coefficient and clustering number are key parameters that affect the performance of FCMT algorithm. The adaptive tuning mechanism improves the robustness across various deployment scenarios as follows:

**Adaptive adjustment of fuzzy coefficients:** Through an adaptive adjustment mechanism, fuzzy coefficients can be dynamically adjusted based on the distribution characteristics of data. When the data distribution is dense and the feature differences are significant, reducing the fuzzy coefficient can enhance the discriminative power of clustering. On the contrary, when there are significant differences in data features, increasing the blur coefficient to accommodate more differences in data features. This adjustment helps achieve the best clustering effect and balance of fuzziness under different data distributions.

**Adaptive adjustment of clustering number:** Based on the size and complexity of the dataset, research is conducted on using density or distance based clustering structure recognition methods to automatically determine the optimal number of clusters. This adaptive adjustment mechanism enables the algorithm to adapt to different data environments and maintain stable clustering and localization performance.

The research introduces an adaptive tuning mechanism that enhances the robustness of the model in various deployment scenarios. This mechanism allows the model to dynamically adjust key parameters based on the characteristics of different scenarios without the need for manual intervention, thereby making the model more adaptable and stable. To demonstrate the adaptability of the model more clearly, the following are some specific deployment scenarios:

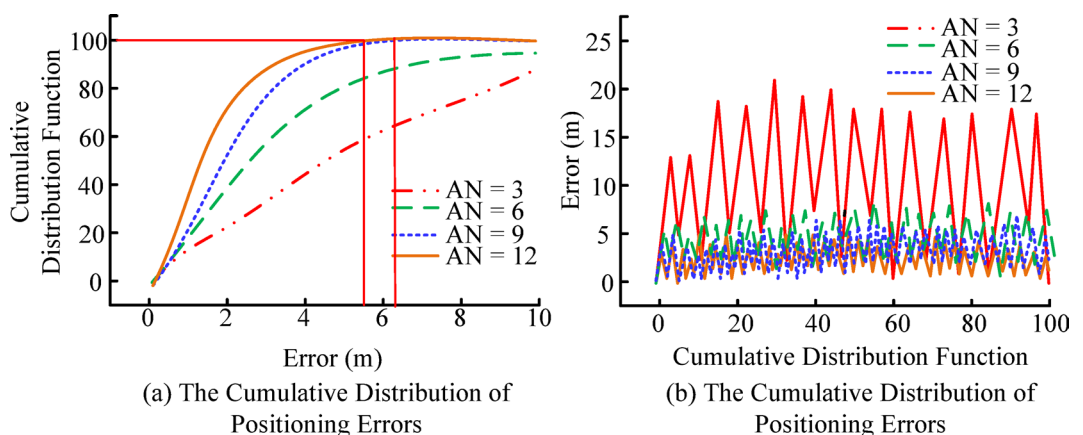
(1) High density node scenario: In densely deployed environments such as industrial automation workshops, the model can optimize fuzzy coefficients to enhance clustering resolution, thereby more accurately locating each node and avoiding signal interference and positioning errors caused by too many nodes. (2) Low density node scenarios: In some environmental monitoring scenarios with wide coverage but few nodes, the model can adjust the number of clusters to ensure accurate division of node distribution in different regions even in sparse data features, thereby improving localization accuracy. (3) Dynamic change scenario: In scenarios where the location and quantity of nodes such as logistics warehouses change over time, the adaptive tuning mechanism of the model can respond to these changes in a timely manner, dynamically adjusting parameters to maintain the accuracy and efficiency of positioning. (4) Complex environmental scenarios: In scenarios with complex signal propagation conditions such as multipath interference and non line of sight propagation (such as inside intelligent buildings), the model can better handle signal uncertainty by optimizing the fuzzy coefficient and clustering quantity, thereby improving the robustness of positioning.

### Analysis of positioning effect of WSN model

The FCMT algorithm showed good clustering performance on both the UCI and MIT Reality datasets. To further explore the positioning accuracy of FCMT-GS-WSN, a hybrid positioning model based on Time of Arrival-Received Signal Strength (TOA-RS) and a distributed positioning model based on Received Signal Strength-Angle-of-Arrival (RSS-AOA) were used as comparative models. TOA-RSS model can combine time arrival information and signal strength information, and RSS-AOA model uses the characteristics of wireless signal strength attenuation with distance to locate, so it is widely used. When using BIM technology to improve positioning accuracy, random deployment allows the network to quickly adapt to changes in the environment, such as node movement or energy consumption, without the need for complex pre-deployment planning. In addition, random deployment provides better coverage, especially in unknown or dynamically changing environments, which is critical for application scenarios that require rapid deployment and high robustness. In practical applications, increasing the number of ANs can improve the accuracy of positioning, but it also increases deployment and computation costs. To investigate the impact of the amount of ANs on the positioning accuracy of FCMT-GS-WSN, different numbers of anchor point systems were tested, and the test outcomes are denoted in Fig. 11.

In Fig. 11 (a), when AN=12, the FCMT-GS-WSN had the most accurate positioning, and the PE of all test points was below 5.41 m. When AN=9, the PE of all test points was below 6.30 m. In Fig. 11 (b), when AN=12, the error fluctuation range of the test point was the smallest, and the highest error was 4.81 m at this time. When AN=9, the error of the test point remained stable and not significant, with a maximum error of 5.41 m. But when AN was less than 6, the error fluctuation of the test point was large, especially when AN=3, the maximum error of the test point reached 22.82 m. From this, under the influence of multidimensional Gaussian model optimization, FCMT-GS-WSN had good localization performance. Although a higher AN can ensure better positioning performance, it will increase deployment costs and resource consumption. Therefore, choosing AN 9 is more suitable. The AN was determined to be 9. By collecting the positioning errors of each model under different communication radii of 25 m to 50 m, the average value is calculated, and the trend of error change with the communication radius is observed. The average PE of each model under different communication radii is shown in Table 4.

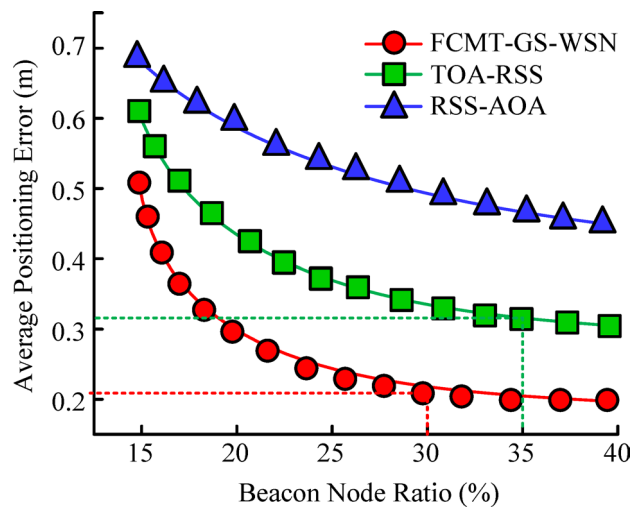
The data in Table 4 shows that as the communication radius increased, the average PEs of FCMT-GS-WSN, TOA-RS, and RSS-AOA all decreased. Due to the larger communication radius, the number of hops in inter node communication was reduced, thereby reducing the cumulative error in multi hop communication. When the communication radius was 25 m, the average PEs of the three models were 0.52 m, 0.63 m, and 0.69 m, respectively. However, when the communication radius increased to 50 m, the errors decreased to 0.21 m, 0.39 m, and 0.42 m, respectively. Among them, FCMT-GS-WSN showed good positioning accuracy under different communication radii. When the communication radius was 40 m, the average PE stabilized at a low level of 0.21 m. The TOA-RS model had a large error in small radii, but as the communication radius increased,



**Fig. 11.** Test point positioning error.

Communication radius (m)	FCMT-GS-WSN	TOA-RSS	RSS-AOA
25	0.52 m	0.63 m	0.69 m
30	0.43 m	0.55 m	0.60 m
35	0.32 m	0.50 m	0.53 m
40	0.21 m	0.46 m	0.48 m
45	0.21 m	0.42 m	0.45 m
50	0.21 m	0.39 m	0.42 m

**Table 4.** The average positioning error of each model under different communication radius.

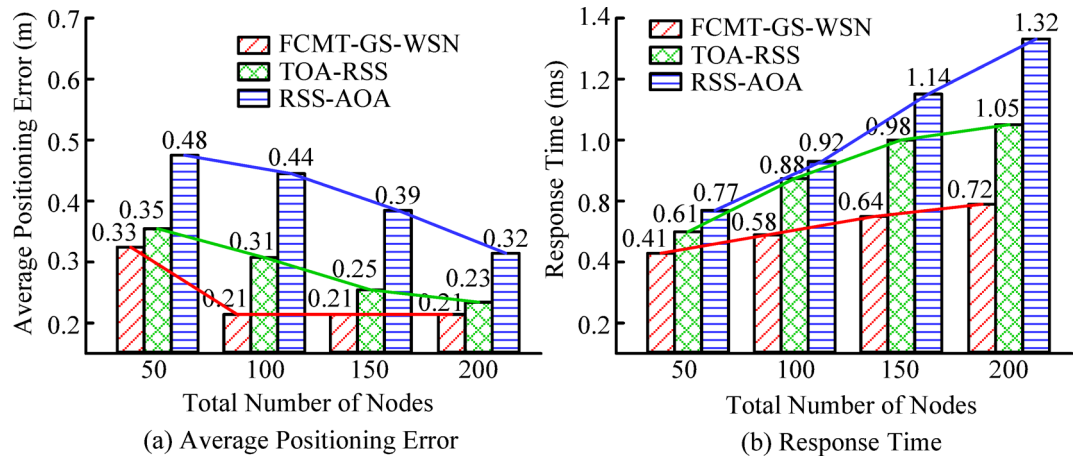


**Fig. 12.** The positioning error of the model under different ratio of beacon nodes.

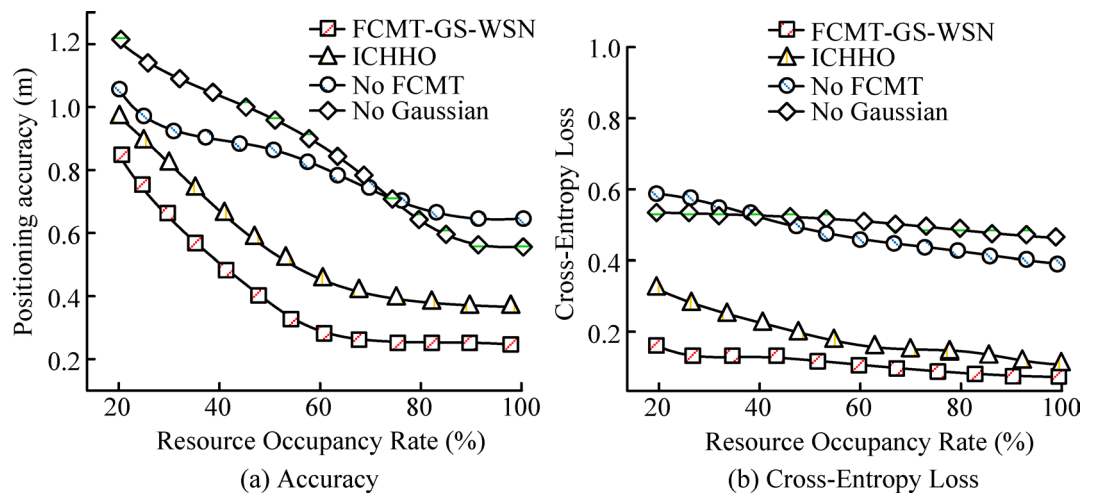
the error decreased significantly. In contrast, the RSS-AOA model showed a relatively small decrease in error as the communication radius increased, indicating limited performance improvement under large radii. From this, FCMT-GS-WSN can perform precise positioning under different communication radius conditions, solving the impact of communication radius on network node positioning. The PEs of each model under different beacon node ratios are denoted in Fig. 12.

According to Fig. 12, as the proportion of beacon nodes in the total number of nodes increased, the average PE of each model showed a decreasing trend. Due to the increase in the amount of beacon nodes, the amount of hops from unknown nodes to beacon nodes decreased, thereby reducing path errors caused by multi hop communication. In addition, FCMT-GS-WSN could obtain more accurate hop count estimation by further refining the communication radius of network nodes. Therefore, when the proportion of beacon nodes reached 30–40%, the average error tended to stabilize at 0.21 m. The TOA-RS model only stabilized at 0.32 m when the proportion of beacon nodes increased to 35%. The average PE of the RSS-AOA model remained unstable even when the proportion of beacon nodes increased to 40%. The FCMT-GS-WSN reduced the average error by 52.38% compared to the TOA-RS model. From this, it can be seen that FCMT-GS-WSN can still ensure positioning accuracy and precision when the proportion of beacon nodes is low, reducing costs and energy consumption. When the beacon node is 35, the average error and positioning response time of each model under different total node numbers are shown in Fig. 13.

In Fig. 13 (a), the average error of each model decreased with the increase of the amount of nodes. This is because increasing the number of nodes can provide more path options, which helps to choose shorter or better transmission paths, reducing signal attenuation and errors during transmission. Among them, FCMT-GS-WSN stabilized at 0.21 when the total number of nodes reached 100, while TOA-RS and RSS-AOA models still showed a downward trend until the total amount of nodes reached 200. In Fig. 13 (b), the response time of each model increased with the increase of the amount of nodes. This is because more nodes mean more measurement data and reference points, which undoubtedly increases the computational complexity of the model. As the amount of nodes increased, the response time of FCMT-GS-WSN grew slowly. When the amount of nodes was 50, the response time was 0.41ms, which was 0.20 ms faster than the second TOA-RSS model. When the amount of nodes increased to 200, the response time of FCMT-GS-WSN was 0.72 ms, which was 0.33 ms faster than the second TOA-RSS model. In summary, FCMT-GS-WSN can achieve high positioning accuracy with a small number of nodes, and can also ensure fast response in scenarios with a large number of nodes, demonstrating versatility in different scenarios. In the practical application of a certain logistics warehouse, the influence of FCMT model and Gaussian model in FCMT-GS-WSN on the whole framework is analyzed and tested. The Improved Chimp



**Fig. 13.** The average error and positioning response time of the model under different total number of nodes.



**Fig. 14.** Positioning accuracy and cross entropy loss of each model under different resource utilization.

Optimization and Harris Hawk Optimization Algorithm (ICHHO) was also used as a comparison model<sup>41</sup>. The test results are shown in Fig. 14.

In Fig. 14 (a), when the resource utilization rate is 62.57%, FCMT-GS-WSN can maintain a relatively high accuracy at 0.28 m while ensuring a relatively low resource utilization rate. In contrast, the ICHHO algorithm's positioning accuracy is 0.41 m. When the FCMT-GS-WSN does not adopt the FCMT algorithm, the positioning accuracy is 0.79 m, and when the Gaussian model is not adopted, the positioning accuracy is 0.84 m. It can be seen that the Gaussian model has a greater influence on the positioning accuracy of FCMT-GS-WSN. The positioning accuracy of FCMT-GS-WSN is improved by 0.13 m compared with ICHHO algorithm. In Fig. 14 (b), when the resource occupancy rate is 62.57%, the cross entropy loss of FCMT-GS-WSN is 0.09, and that of ICHHO algorithm is 0.13. It can be seen that FCMT-GS-WSN can predict the position more accurately. In the actual test, FCMT-GS-WSN can accurately locate the logistics location of the logistics warehouse under the condition of low resource utilization.

### Computational complexity and energy analysis

In terms of computational complexity, the time complexity of FCMT algorithm mainly depends on the fuzzy coefficient and the number of feature points. When the blur coefficient is 2 and the number of feature points is 100, the time complexity is  $O(100 \times \log \times 100)$ . This indicates that as the number of feature points increases, the processing time will increase logarithmically, making it a relatively efficient algorithm. In addition, the computational complexity of the multidimensional Gaussian model is  $O(50 \times 50)$ , which is relatively stable in processing time when the sample size is 50. Overall, the FCMT-GS-WSN model has low computational complexity and is suitable for application in resource constrained WSN environments.

In terms of energy analysis, node energy consumption mainly comes from factors such as sampling frequency, node density, and communication radius. When the sampling frequency is 100 Hz, the node density is 0.2 nodes per square meter, and the communication radius is 30 m, the average energy consumption of nodes is 0.42 joules

per hour. This indicates that at a reasonable sampling frequency and node density, the energy consumption of nodes is at a lower level, which helps to extend the lifespan of the network. In addition, by optimizing the sampling box and fuzzy inference algorithm, the energy consumption of nodes can be further reduced. When the communication radius increases to 40 m, energy consumption can be reduced to 0.35 joules per hour, indicating that increasing the communication radius helps to reduce energy consumption, but the relationship between communication costs and energy consumption needs to be balanced.

## Conclusion

This study mainly focuses on the positioning technology in WSNs. To address the challenges of traditional positioning methods in terms of accuracy and efficiency, a WSN positioning technology based on improved sampling boxes and fuzzy reasoning was proposed. The study adopted a fuzzy clustering algorithm based on time series optimization and combined it with a multidimensional Gaussian model to optimize the sampling box, to raise the accuracy and efficiency of localization. The research results indicated that the FCMT algorithm exhibited good clustering performance and robustness under different fuzzy coefficients and weight allocation indices. Tests on the UCI and MIT Reality datasets showed that the clustering performance of the FCMT algorithm was superior to the comparison algorithms, with its RI remaining stable at 0.89 when the weight allocation index was 5, significantly higher than other algorithms. In addition, the average PE of FCMT-GS-WSN was lower than other models under different communication radii and beacon node ratios. When the proportion of beacon nodes reached 40%, the average error of FCMT-GS-WSN stabilized at 0.21 m, which was 52.38% lower than the TOA-RS model. As the amount of nodes increased, the response time of FCMT-GS-WSN grew slowly. When the number of nodes was 50 and 200, the response times of FCMT-GS-WSN were 0.41ms and 0.72ms, respectively, which were 0.20ms and 0.33ms faster than the second TOA-RS model. The positioning accuracy of FCMT-GS-WSN when the resource usage is 62.57% is improved by 0.13 m compared with ICHHO algorithm. In summary, the FCMT-GS-WSN model can maintain fast response in scenarios with a large number of nodes while ensuring positioning accuracy, demonstrating versatility and efficiency in different scenarios. Although research has made progress in improving the positioning accuracy of WSN, the performance of the FCMT-GS-WSN model still needs to be further verified under extreme environmental conditions. For instance, in harsh environments such as high temperatures, high humidity, and strong electromagnetic interference, positioning accuracy may be affected. In the future, the FCMT-GS-WSN model will be applied to various extreme environments, and a large number of experiments will be conducted to collect data and analyze the specific influence of different environmental factors on the positioning accuracy, thereby dynamically optimizing the algorithm parameters. In addition to the existing application scenarios, we also plan to extend the FCMT-GS-WSN model to more fields. For example, irrigation equipment used to monitor the growth environment of crops and achieve precise positioning. In intelligent transportation, it is used for high-precision positioning of vehicles and traffic flow monitoring, etc.

## Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 24 April 2025; Accepted: 8 July 2025

Published online: 18 July 2025

## References

- Zagrouba, R. & Kardi, A. Comparative study of energy efficient routing techniques in wireless sensor networks, *Inf. Jan* **12** (1), 42–58. <https://doi.org/10.3390/info12010042> (2021).
- Lopez-Ardao, J. C., Rodriguez-Rubio, R. F., Suarez-Gonzalez, A., Rodriguez-Perez, M. & Sousa-Vieira, M. E. Current trends on green wireless sensor networks, *Sens. Jul* **21** (13), 4281–4294. <https://doi.org/10.3390/s21141407> (2021).
- Kanoun, O. et al. Jan., Energy-aware system design for autonomous wireless sensor nodes: A comprehensive review, *Sens.*, **21**, 2, pp. 548–561, (2021). <https://doi.org/10.3390/s21020548>
- Sahoo, S. K. et al. Intelligent trust-based utility and reusability model: enhanced security using unmanned aerial vehicles on sensor nodes, *Appl. Sci.*, vol. 12, no. 3, pp. 1317–1334, Feb. (2022). <https://doi.org/10.3390/app12031317>
- Bhosle, K. & Musande, V. Evaluation of deep learning CNN model for recognition of devanagari digit. *Artif. Intell. Appl.* **1** (2), 114–118. <https://doi.org/10.47852/bonviewAIA3202441> (Feb. 2023).
- Mani, R., Rios-Navarro, A., Sevillano-Ramos, J. L. & Liouane, N. Improved 3D localization algorithm for large scale wireless sensor networks. *Wirel. Netw.* **30** (6), 5503–5518. <https://doi.org/10.1007/s11276-023-03265-0> (Jun. 2024).
- Lou, H. et al. DC-YOLOv8: Small-size object detection algorithm based on camera sensor, *Electron. May* **12** (10), 2323–2335. <https://doi.org/10.3390/electronics12102323> (2023).
- Liu, G. et al. Feb., An enhanced intrusion detection model based on improved kNN in wsns, *Sens.*, **22**, 4, pp. 1407–1425, (2022). <https://doi.org/10.3390/s22041407>
- Safaldin, M., Otair, M. & Abualgiah, L. Improved binary Gray Wolf optimizer and SVM for intrusion detection system in wireless sensor networks. *J. Ambient Intell. Humaniz. Comput.* **12** (1), 1559–1576. <https://doi.org/10.1007/s12652-020-02228-z> (Jan. 2021).
- Cao, C., Tang, Y., Huang, D. Y., Gan, W. & Zhang, C. IIBE: An improved identity-based encryption algorithm for wsn security, *Secur. Commun. Netw.*, vol. no. 1, p. 8527068, Sep. 2021, (2021). <https://doi.org/10.1155/2021/8527068>
- Kanwar, V. & Kumar, A. DV-Hop localization methods for displaced sensor nodes in wireless sensor network using PSO, *Wirel. Netw.*, vol. 27, no. 1, pp. 91–102, Jan. (2021). <https://doi.org/10.1007/s11276-020-02417-2>
- Kathiroli, P. & Selvadurai, K. Energy efficient cluster head selection using improved sparrow search algorithm in wireless sensor networks. *J. King Saud Univ. - Comput. Inf. Sci.* **34** (10), 8564–8575. <https://doi.org/10.1016/j.jksuci.2022.06.007> (Oct. 2022).
- Lakshmana, K. et al. Improved metaheuristic-driven energy-aware cluster-based routing scheme for IoT-assisted wireless sensor networks, sustainability. *Jul* **14** (13), 7712–7724. <https://doi.org/10.3390/su14137712> (2022).
- Mohan, P. et al. Improved metaheuristics-based clustering with multihop routing protocol for underwater wireless sensor networks, *Sens. Feb* **22** (4), 1618–1634. <https://doi.org/10.3390/s22041618> (2022).

15. Roy, S., Mazumdar, N. & Pamula, R. An optimal mobile sink sojourn location discovery approach for the energy-constrained and delay-sensitive wireless sensor network. *J. Ambient Intell. Humaniz. Comput.* **12** (12), 10837–10864. <https://doi.org/10.1007/s12652-021-03439-9> (Dec. 2021).
16. Ahmadi, A. et al. Wireless network deployment survey, *IEEE Radio Wirel. Symp.*, pp. 143–146, Feb. (2024). <https://doi.org/10.1109/RWS56914.2024.10438583>
17. Chambon, A., Sahli, A., Rachedi, A. & Mebarki, A. Optimizing IoT networks deployment under connectivity constraint for dynamic digital twin, 2023 IEEE int. Conf. *Metaverse Comput. Netw. Appl.* 474–480. <https://doi.org/10.1109/MetaCom57706.2023.00088> (Oct. 2023).
18. Zaimen, K., Moalic, L., Brahmia, M. E. A., Abouaissa, A. & Idoumghar, L. A hybrid binary grey Wolf optimiser for WSN deployment in indoor environments based on BIM database. *IEEE Wirel. Commun. Netw. Conf.* 1–6. <https://doi.org/10.1109/WCNC57260.2024.10570720> (Jul. 2024).
19. Hada, R. P. S., Aggarwal, U. & Srivastava, A. A Study and Analysis of a New Hybrid Approach for Localization in Wireless Sensor Networks, *J. Web Eng.*, vol. 22, no. 2, pp. 279–302, March, (2023). <https://doi.org/10.13052/jwe1540-9589.2224>
20. Nguyen, T. T., Dao, T. K., Ngo, T. G. & Nguyen, T. D. Node WSN localisation based on adaptive crossover-mutation differential evolution, *Int. J. Sensor Netw.*, vol. 44, no. 1, pp. 1–22, January, (2024). <https://doi.org/10.1504/IJSNET.2024.136339>
21. Gulati, K. et al. A review paper on wireless sensor network techniques in Internet of Things (IoT), *Mater. Today: Proc.*, vol. 51, no. 1, pp. 161–165, May. (2022). <https://doi.org/10.1016/j.matpr.2021.12.448>
22. Lanzolla, A. & Spadavecchia, M. Wireless sensor networks for environmental monitoring, *Sens.*, vol. 21, no. 4, pp. 1172–1184, Feb. (2021). <https://doi.org/10.3390/s21041172>
23. Yan, R., Yu, Y. & Qiu, D. Emotion-enhanced classification based on fuzzy reasoning. *Int. J. Mach. Learn. Cybern.* **13** (3), 839–850. <https://doi.org/10.1007/s13042-021-01156-9> (Mar. 2022).
24. Jiang, W., Zhou, K. Q., Sarkheyli-Hägele, A. & Zain, A. M. Modeling, reasoning, and application of fuzzy petri net model: a survey. *Artif. Intell. Rev.* **55** (8), 6567–6605. <https://doi.org/10.1007/s10462-021-09856-7> (Dec. 2022).
25. Zhang, X., Wang, M., Bedregal, B., Li, M. & Lian, R. Semi-overlap functions and novel fuzzy reasoning algorithms with applications. *Inf. Sci.* **614** (1), 104–122. <https://doi.org/10.1016/j.ins.2021.06.044> (Jan. 2022).
26. Masini, R. P., Medeiros, M. C. & Mendes, E. F. Machine learning advances for time series forecasting. *J. Econ. Surv.* **37** (1), 76–111. <https://doi.org/10.1111/joes.12456> (Feb. 2023).
27. Blázquez-García, A., Conde, A., Mori, U. & Herrera, F. A review on outlier/anomaly detection in time series data. *ACM Comput. Surv. (CSUR)*. **54** (3), 1–33. <https://doi.org/10.1145/3453586> (Aug. 2021).
28. Zhou, H. et al. Informer: Beyond efficient transformer for long sequence time-series forecasting, *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 12, pp. 11106–11115, May. (2021). <https://doi.org/10.1609/aaai.v35i12.18045>
29. Schmidl, S., Wenig, P. & Papenbrock, T. Anomaly detection in time series: a comprehensive evaluation, *Proc. VLDB Endow.*, vol. 15, no. 9, pp. 1779–1797, Jul. (2022). <https://doi.org/10.14778/3539038.3539040>
30. Tang, Y., Pan, Z., Pedrycz, W., Ren, F. & Song, X. Based kernel fuzzy clustering with weight information granules. *IEEE Trans. Emerg. Top. Comput. Intell.* **7** (2), 342–356. <https://doi.org/10.1109/TETCI.2022.3187379> (Jun. 2022).
31. Pan, Y., Li, Q., Liang, H. & Lam, H. K. A novel mixed control approach for fuzzy systems via membership functions online learning policy. *IEEE Trans. Fuzzy Syst.* **30** (9), 3812–3822. <https://doi.org/10.1109/TFUZZ.2022.3202463> (Sep. 2022).
32. Xing, Y., Rappaport, T. S. & Ghosh, A. Millimeter wave and sub-THz indoor radio propagation channel measurements, models, and comparisons in an office environment. *IEEE Commun. Lett.* **25** (10), 3151–3155. <https://doi.org/10.1109/LCOMM.2021.3111832> (Oct. 2021).
33. Li, Y. et al. Enhanced RSS-based UAV localization via trajectory and multi-base stations. *IEEE Commun. Lett.* **25** (6), 1881–1885. <https://doi.org/10.1109/LCOMM.2021.3074986> (Jun. 2021).
34. Dong, Y., Arslan, T. & Yang, Y. Real-time NLOS/LOS identification for smartphone-based indoor positioning systems using WiFi RTT and RSS. *IEEE Sens. J.* **22** (6), 5199–5209. <https://doi.org/10.1109/JSEN.2021.3053629> (Mar. 2021).
35. Zou, Y. & Liu, H. RSS-based target localization with unknown model parameters and sensor position errors. *IEEE Trans. Veh. Technol.* **70**, 6969–6982. <https://doi.org/10.1109/TVT.2021.3067194> (Jul. 2021).
36. Qiu, Y. L., Ma, L. K. & Priyadarshi, R. Deep learning challenges and prospects in wireless sensor network deployment, *Arch. Comput. Methods Eng.*, vol. 31, no. 6, pp. 3231–3254, March, (2024). <https://doi.org/10.1007/s11831-024-10079-6>
37. Priyadarshi, R., Soni, S. K. & Nath, V. Energy efficient cluster head formation in wireless sensor network. *Microsyst. Technol.* **24** (1), 4775–4784. <https://doi.org/10.1007/s00542-018-3873-7> (April, 2018).
38. Priyadarshi, R., Gupta, B. & Anurag, A. Deployment techniques in wireless sensor networks: a survey, classification, challenges, and future research issues, *J. Supercomput.*, vol. 76, no. 1, pp. 7333–7373, January, (2020). <https://doi.org/10.1007/s11227-020-03166-5>
39. Bahadur, D. K. J. & Lakshmanan, L. A novel method for optimizing energy consumption in wireless sensor network using genetic algorithm, *Microprocess. Microsyst.*, vol. 96, no. 1, pp. 104749–104750, February, (2023). <https://doi.org/10.1016/j.micpro.2022.104749>
40. Kooshari, A. et al. An optimization method in wireless sensor network routing and IoT with water strider algorithm and ant colony optimization algorithm, *Evolut. Intell.*, vol. 17, no. 3, pp. 1527–1545, April, (2024). <https://doi.org/10.1007/s12065-023-00847-x>
41. R. JIA and H. ZHANG. Improved population intelligence algorithm for wireless sensor network coverage optimization. *J. Southwest. Univ. Nat. Sci. Ed.*, **46**, 1, pp. 155–166, January, (2024). <https://doi.org/10.13718/j.cnki.xdzk.2024.01.013>

## Author contributions

N.Y.D. processed the numerical attribute linear programming of communication big data, and the mutual information feature quantity of communication big data numerical attribute was extracted by the cloud extended distributed feature fitting method. L.Z. and W.J.P. Combined with fuzzy C-means clustering and linear regression analysis, the statistical analysis of big data numerical attribute feature information was carried out, and the associated attribute sample set of communication big data numerical attribute cloud grid distribution was constructed. N.Y.D. and Y.C. did the experiments, recorded data, and created manuscripts. All authors read and approved the final manuscript.

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

Correspondence and requests for materials should be addressed to L.Z.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025