Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105
Volume **9**, Number **4**, November **2024**

WSN Data Fusion Algorithm Based on Radial Cluster and Elephant Swarm Neural Networks

Jia-Wei Liu

School of Computer Science and Technology
Henan Polytechnic University, Jiaozuo 454000, China
18439652738@163.com

Wei-Peng An*

School of Software
Henan Polytechnic University, Jiaozuo 454000, China
212209020082@home.hpu.edu.cn

*Corresponding author: Wei-Peng An

Received December 1, 2023, revised March 17, 2024, accepted June 24, 2024.

Abstract. *To minimize the propagation of redundant data in wireless sensor networks, conserve energy, and extend network lifespan, we propose an algorithm (R-IEHOBP) that combines radial clustering and an elephant swarm neural network. Initially, we suggest employing the radial clustering routing protocol to address the issue of uneven energy consumption in nodes caused by the majority of current data fusion algorithms based on the LEACH clustering routing protocol. In the radial clustering protocol, we first divide candidate cluster heads into clusters of different sizes, and then we screen the cluster heads based on the number of required cluster head nodes in the sub-cluster, the remaining energy of the candidate nodes, distance from the center of gravity, and the number of nearby one-hop nodes. Secondly, we improve the elephant swarm optimization algorithm by introducing chaotic sequences, elite strategies, adversarial learning, and guided substitution principles. Next, we use the elephant group algorithm to select the initial matrices of weights and thresholds for the neural network to avoid the impact of unsuitable initial parameters on the final convergence speed and output accuracy. Addressing issues in single-stage neural network data fusion, such as incomplete data processing and large errors in the fused data, this paper proposes using a second-stage neural network to fuse the initial data received by the cluster head node and output it to the aggregation node. Finally, we establish a WSN data fusion model based on the radial clustering routing protocol by combining the radial clustering structure of wireless sensor networks with the neural network structure. Simulation experiment results demonstrate that the algorithm exhibits better redundant data rejection ability compared to similar algorithms, with lower energy consumption and higher data fusion accuracy.*
Keywords:WSN, Radial clustering, Elephant herd optimization algorithm, Neural networks, Data fusion

1. **Introduction.** WSN (wireless sensor networks) is a self-organized network system consisting of multiple sensor nodes with multi-functions such as communication, data acquisition, and processing [1]. Sensors are powered by dry batteries carried by themselves, and in most application environments, the batteries cannot be replaced. Therefore, the efficient utilization of node energy and the reduction of the average node's network energy consumption have become hot issues currently studied by various scholars [2]. To ensure

the comprehensiveness and real-time nature of the monitoring data, the staff tends to construct the monitoring and early warning network by throwing sensor nodes in high density. A large amount of redundant data in the system mainly comes from two aspects, one of which is the high density of sensors in some areas, which leads to the repeated collection of the same location data by different sensors. The other is the invalid transmission data generated by the negligible fluctuation of some data during uninterrupted data collection [3]. Theoretical analysis demonstrates that the most energy-intensive stage in the operation of wireless sensors is the data transmission phase. Consequently, the pivotal aspect for extending the longevity of nodes and the overall operational lifespan of the network lies in the effective reduction of redundant data transmission within the system [4].

1.1. **Related work.** The following content provides a concise literature review of the method proposed in this paper. There is an essential difference between data fusion and classical signal processing methods, data fusion deals with more complex forms of multi-sensor information and is possible at different information levels, where each level represents a different degree of fusion process on the data, and this information abstraction levels contain data level, feature level and decision level. The corresponding fusion methods are also mainly data level, feature level, and decision level fusion [5]. The research goal of this paper is feature-level data fusion, so the rest of the data fusion levels will not be covered. Feature-level data fusion belongs to the intermediate level of data fusion, and the main advantage is that it compresses the original data and reduces the propagation of redundant or interfering data, which is advantageous in real-time data processing and has high accuracy. The methods usually used in feature-level fusion are cluster analysis, Bayesian estimation, information entropy, weighted average, voting and neural network methods, etc. Since neural network technology has the same characteristics as data fusion technology, a large number of scholars have carried out research on WSN data fusion based on neural network technology [6, 7] .

Wang et al. [8] introduced a BP neural network data fusion model based on the TEEN protocol. The model initially filters redundant data by setting the TEEN threshold and then conducts data fusion with the neural network. Results indicate a reduction in communication and energy consumption in WSNs, enhancing data collection efficiency. Nevertheless, challenges include uneven energy utilization among nodes and an unstable model convergence speed. Wang et al. [9] proposed a hybrid approach, combining rough sets and neural networks. Rough sets are employed to simplify network inputs, improving the training speed. However, the implementation process is influenced by the accuracy of the original decision table. Ayhan et al. [10] suggested optimizing neural network weights and structure using genetic algorithms, creating a data fusion model. While extending the network's lifespan to some extent, this approach faces limitations in processing scale and stability. Sun et al. [11] presented a wireless sensor network data fusion algorithm based on BP neural networks. The algorithm, integrated with the LEACH routing protocol, utilizes neural networks for feature extraction from data collected by cluster members. The aggregated data is then transmitted to the convergence node, effectively reducing data transmission traffic and node transmission energy. Notably, the sensitivity of initial values in BP neural networks introduces a challenge, potentially leading to locally optimal solutions.

The inherent sensitivity of neural networks to initial values poses a challenge; an inaccurately chosen threshold matrix initial value can significantly impede the model's iteration speed. Intelligent optimization algorithms, a category designed for swift convergence to optimal values, serve as a remedy for this neural network shortcoming. Scholars have ingeniously proposed employing these intelligent algorithms to optimize the selection of

weights and initial matrices for neural networks [12]. Cao et al. [13] proposed a WSN data fusion algorithm based on the improved Gray Wolf algorithm optimized BP neural network, which improves the data fusion accuracy under different datasets and reduces the node energy consumption to some extent. Yu et al. [14] introduced a data fusion algorithm that combines an improved ant colony approach with a BP neural network. This innovative method employs an enhanced ant colony algorithm to optimize the performance of the BP neural network, specifically tailored for deep well scenarios. Koutini et al. [15] proposed a neural network data fusion algorithm that was carefully tuned by applying the Tennessee Whisker algorithm. This optimization not only improves the accuracy of data fusion but also speeds up the convergence of the network. However, a noteworthy limitation is that the algorithm neglects the delicate balance between local optimization and global search, and fails to address the problem that particle swarms are prone to fall into local optima.

1.2. **Main contribution.** Faced with a series of problems encountered by various intelligent optimization algorithms in fusing BP neural network models, such as low convergence of the algorithms, the search process is very easy to fall into the local optimal solution, there is still a large amount of redundant data in the fusion process, as well as the high complexity of the actual operation process, this paper proposes the radial distribution-based hierarchical cluster neural network data fusion algorithm (R-IEHOBP). The main contributions of our work are as follows:

(1) We combine the radial clustering structure of wireless sensor networks with the neural network structure to establish a radial clustering-based data fusion model for wireless sensor networks [16, 17]. The problem of uneven energy utilization of nodes caused by traditional neural network-like data fusion algorithms based on LEACH clustering routing protocol is solved to avoid the generation of energy voids, which can effectively extend the network lifetime.

(2) We propose to improve the classical elephant swarm optimization algorithm using chaotic mapping, guided substitution principle, and elite strategy. It enhances the diversity of the initial population of the elephant swarm, enhances the global search ability of the elephant swarm algorithm, prevents the elephant swarm algorithm from falling into local optimal solutions, and speeds up the iteration speed of the algorithm.

(3) We use the improved elephant swarm optimization algorithm to select the initial weights and threshold matrix of the neural network, which avoids the adverse effects of wrong initial values of the biblical network on the iterative process [18].

(4) We advocate a two-step approach in the data fusion process, employing a two-layer neural network structure. The initial layer of the neural network undertakes data preprocessing, forwarding the processed data to the second layer for fusion operations. This design departs from a single-layer neural network data fusion model, addressing issues related to inadequate data processing and low fusion accuracy through optimization [19].

2. **Radial Cluster Protocol.** In this paper, a radial distribution-based cluster head selection protocol (RACH) is used. Its advantage lies in abandoning the previous way of directly selecting cluster heads in the whole network, dividing the network into arches of different sizes, determining the number of required cluster head nodes in the arched according to the arc length and radius of each arched, and then scattering the cluster head campaigning process in each arched. RACH not only reduces the energy consumption during cluster head campaigning but also solves the problem of energy wastage caused by too little cluster head node selection when the cluster head consumes a lot of energy

and is prone to node death when too many cluster head nodes are selected. The RACH structure diagram is shown in Figure 1.

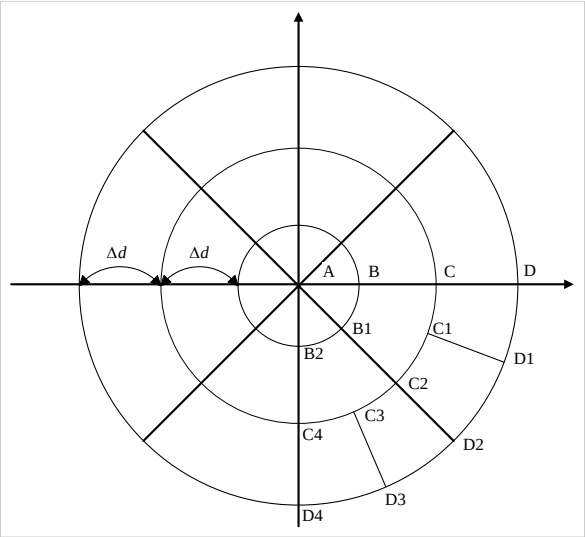

FIGURE 1. RACH structure diagram

In the cluster head selection phase, the number of cluster heads needed in the region is first determined according to Equation (1). and then the node weights are calculated according to the number of neighboring junctions, the remaining energy of the node, and the distance from the center of gravity of the region for each candidate cluster head, and the cluster head nodes are selected according to the weights, as shown in Equation (2).

$$S_i = Trunc(P_i/R_i \times 4) + 1 \tag{1}$$

$$\begin{cases} D_{s,i} = \sqrt{(X_s - X_i)^2 + (Y_s - Y_i)^2} \\ w_{(s)} = Er(s) + N(s) - D_{s,i}/100 \end{cases} \tag{2}$$

Where P_i is the arc length of the ith sector, R_i denotes the radius of the ith sector, S_i represents the optimal number of cluster heads needed in the ith arch, represents the distance of node s from the center of gravity of the region in which it is located, $w_{(s)}$ represents the weight of node s, $D_{s,i}$ is the residual energy of the node, and $Er(s)$ is the number of one-hop nodes of the node in the same region.

3. Improvement of EHO.

3.1. **Improved initialization algorithm.** Owing to the pronounced uncertainty associated with the randomly generated initial population in the standard EHO algorithm, it can readily impart adverse effects on the iterative process. Consequently, this article presents a novel approach by introducing an initial value generation algorithm rooted in chaotic mapping and Opposition-based Learning (OBL) [20]. Chaotic mapping has the characteristics of pseudo-randomness, ergodicity, and unpredictability, which can be utilized to map the randomly generated initial positions in the algorithm into the chaotic space to make the initialized population positions more uniform, thus increasing the population diversity. Numerous studies have shown that Tent mapping has significant advantages in terms of traversal consistency, so we use Tent mapping to map the initial value positions of object groups. Firstly, the classical elephant swarm algorithm initializes its population

through a random uniform distribution that is shown in Equation (3), subsequently, we apply the chaotic mapping to these initial values according to Equation (4). Then the opposite solution of the initial elephant group after chaotic mapping is generated by OBL, as shown in Equation (5). Eventually, the better-adapted individuals are selected to join the initial population.

$$x_{ci,j} = lb_j + (ub_j - lb_j) \times rand \tag{3}$$

$$x_{ci,j+1} = \begin{cases} 2x_{ci,j} 0 \leq x_{ci,j} \leq 0.5 \\ 2(1 - x_{ci,j}) 0.5 < x_{ci,j} \leq 1 \end{cases} \tag{4}$$

$$x'_{ci,j} = (ub_j - lb_j) - x_{ci,j} \tag{5}$$

Where $x_{ci,j}$ represents the generation position of the ci th elephant in the jth dimension of the randomly generated, and $x_{ci,j+1}$ Represents the value of $x_{ci,j}$ after Tent mapping, and $x'_{ci,j}$ stands for the opposite value of $x_{ci,j}$, and ub_j and lb_j denote the upper and lower bounds in the space, respectively, where rand $\in$ [0,1].

3.2. **Optimizing the outlier approach.** The classical elephant swarm algorithm evicts the elephant with the lowest fitness in each separation step, which not only affects the diversity of the elephant swarm but also easily leads to the overall search results falling into local optima. Certain scholars suggest incorporating randomly generated new elephant individuals to substitute for the least fit individuals. However, this may not yield effective optimization results during the algorithm's iteration and could potentially cause the algorithm to iterate in an incorrect direction. This article proposes a guiding replacement strategy related to the position of the worst common elephant and leader, replacing the classical elephant outlier algorithm. As shown in Equation (6).

$$x_{new} = \begin{cases} x_{best} + r1(x_{best} - x_{worst}) + r_2(x_1 - x_2), r > 0.5 \\ x_{worst} + r1(x_{best} - x_{worst}) + r_2(x_1 - x_2), r < 0.5 \end{cases} \tag{6}$$

Where x_{new} Represents newly generated replacement elephant individuals, r_1 and r_2 are the random values that between 0 and 1, x_1 and x_2 are randomly removed individual elephants in this clade and $x_1 \neq x_2$.

3.3. **Improving random parameter sequences.** Chaotic sequences are a type of stochastic process that maintains the unpredictability of random numbers while also possessing non-periodicity and other characteristics. The population initialization, search, update, and crossover processes involving chaotic sequences can achieve better results than random sequences. Tuba et al. compared the effect of Circle mapping and Sinusoidal mapping in the hierarchical swarm optimization algorithm through simulation experiments, and the experimental results show that the hierarchical swarm optimization algorithm with Circle mapping has the optimal results under the same initial environment [21]. In order to avoid the population iteration falling into the local optimum and to speed up the algorithm to search for the global optimum solution, Circle mapping is introduced in this paper to replace the pseudo-random value generator, that shows in Equation (7).

$$s_{k+1} = \mathrm{mod}\left(s_k + a - \frac{b}{2\pi}\sin(2\pi s_k), 1\right) \tag{7}$$

Where s_{k+1} represents a newly generated chaotic sequence, used to replace the random value generator in the text, with values between (0,1), when $a = 0.2$ and $b = 0.5$, the generated chaotic sequence lies within (0,1).

3.4. **Introducing Elite Strategies to Optimize EHO Leader Update Operations.** In order to solve the problem of leader reverse updating in the standard elephant swarm algorithm, this paper introduces an elite strategy in the leader updating process. By comparing the fitness of the current herd leader $X_{ci,j}^{t}$ with the new leader $X_{ci,j}^{t+1}$ generated in subsequent iterations, the one with the highest fitness is selected as the new leader. The elite strategy can be abstracted as shown in Equation (8).

$$X_{ci,j}^{t+1} = \begin{cases} X_{ci,j}^{t+1} if fit(X_{ci,j}^{t+1}) < fit(X_{ci,j}^{t}) \\ X_{ci,j}^{t} if fit(X_{ci,j}^{t+1}) \geq fit(X_{ci,j}^{t}) \end{cases} \tag{8}$$

Where $X_{ci,j}^{t}$ represents the current herd leader, and $X_{ci,j}^{t+1}$ represents new leaders created during subsequent iterations, represents fitness calculation function.

4. Data Fusion Core Ideas.

4.1. **WSN model structure.** The WSN model in this paper adopts a two-stage neural network structure. The first-level neural network is mainly applied to the preliminary processing of monitoring data in the cluster nodes, which includes data normalization and extraction of data feature values. The second-level neural network is located in the cluster head node, which is mainly responsible for processing the normalized feature data extracted by the first-level neural network. After the data is processed by the implicit layer, the cluster head node delivers the fused data to the base station. The structure of the secondary neural network is shown in Figure 2.

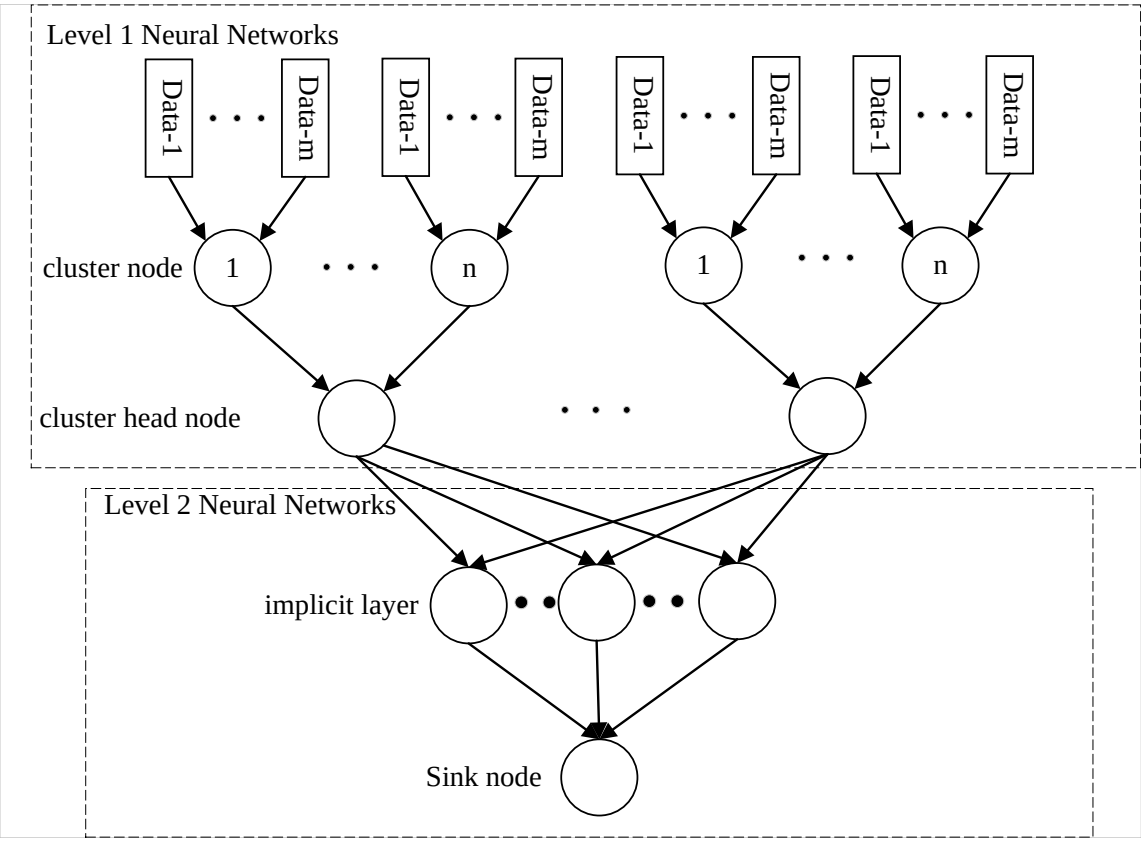

FIGURE 2. Two-stage neural network structure diagram

4.2. **Redundant data elimination.** The main sources of redundant data in the system are data collected repeatedly due to excessive sensor density, and data that fluctuates less than the monitoring error. So in this paper, redundant data elimination is carried out in two parts, firstly the intra-cluster nodes perform preliminary culling of the received data and pass the preliminary culled feature data to the cluster head node for further processing of the data. The experiment proves that the secondary processing of the collected data can significantly reduce the redundant data in the system. The process of data normalization is to calculate the mean value of the current data group by Equation (9), then the mean value is brought into Equation (10) to solve the standard deviation of the data group, and finally the two parameters are brought into Equation (11) to normalize the data.

$$\overline{D} = \frac{1}{n}\sum_{i=1}^{n} d(x_i) \tag{9}$$

$$std(D) = \sqrt{\frac{1}{(n-1)}\sum_{i=1}^{n}\left(d(x_1) - \overline{D}\right)^2} \tag{10}$$

$$G_i = N(d(x_i)) = \frac{d(x_i)}{std(D)} \tag{11}$$

where $d(x_1)$ denotes the set of data to be normalized, and $\overline{D}$ denotes the mean value of the set of data, $std(D)$ denotes the standard deviation of the data for the group, G_i represents the final result after normalization of the data. In this paper, we extract eigenvalues by eliminating combinations of data with fluctuations within the permissible range. Specifically, we do not process data that exhibits fluctuations between the pre-set values θ. In a one-level neural network, the normalized data is passed to the cluster head node through the hidden layer, determine whether to continue processing the data or discard it by comparing the $|d(x_{i+1}) - d(x_i)|$ difference to θ. If $d(\Delta x_i) > \theta$, the data fluctuates beyond the permissible range, and the data is involved in the next processing as characterization data. The feature data extraction method can be abstracted as shown in Equation (12).

$$f(d(x_i), \theta) = \begin{cases} d(\Delta x_i) = |d(x_{i+1}) - d(x_i)| \\ d(x_i), \qquad d(\Delta x_i) < \theta \\ d(x_{i+1}), \qquad d(\Delta x_i) > \theta \end{cases} \tag{12}$$

Where $f(d(x_i), \theta)$ denotes the feature data extraction function, $d(x_i)$ and $d(x_i + 1)$ represents the continuous data passed by the nodes in the cluster, and $d(\Delta x_i)$ represents the absolute value of the difference between two consecutive pieces of data, then θ is the threshold set at the hidden layer node of the first neural network layer.

4.3. **Data fusion.** The purpose of data fusion is to take the monitoring data collected by the WSN and output a single packet that can represent all the current information through the optimization of the neural network implicit layer function. Since the energy consumed for transferring information between sensors is directly proportional to the transmission distance and the size of the information packet, this approach reduces the amount of energy consumed by the cluster head nodes for transferring data. The data fusion process occurs in the second level of neural network structure. The normalized feature data extracted above are used as input values and handed over to the optimization function in the input-implicit layer for further processing. The processed data is then further fused as an input value by the optimization function in the implied-output layer to obtain the

final fused data, which is passed to the convergence node. The optimization function in the input-implicit layer is shown in Equation (13), and The optimization function in the implicit layer-output layer is shown in Equation (14).

$$\begin{cases} H_j = f(\sum_{i=1}^{m} w_{ij} R_i - a_j), j = 1, 2..., p \\ f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \end{cases} \tag{13}$$

$$T_k = \sum_{j=1}^{p} H_j w_{jk} - b_k, k = 1, 2, ..., n \tag{14}$$

Where $f(x)$ is the Tanh excitation function (hyperbolic tangent function), H_j and T_k represent the value output from the implicit layer and the value output at the end of the forward data transfer process, respectively. w_{ij} and w_{jk} are the randomly given link weights, a_j and b_k are the given thresholds at the implicit and output layers, respectively.

4.4. **Data Fusion Steps and Flowchart.** In this paper, a WSN data fusion algorithm based on radial clustering and IEHOBP is proposed, and the specific implementation steps are as follows:

(1) Sensor nodes in the WSN exchange identity information with each other.

(2) Divide and determine the number of sensors required for each region according to Equation (1), and finally calculate the fitness of each node with Equation (2) and select the cluster head node step by step.

(3) The nodes in the cluster normalize the data listened to during the working period according to Equation (9, 10, 11), and then the eigenvalues of the normalized data are extracted by Equation (12).

(4) The intra-cluster node passes the current eigenvalue data collected by various sensors to the cluster head node.

(5) The cluster head node will send the output fused data to the aggregation node after optimizing the received data as an input to the neural network by the implicit layer. The implicit layer function is optimized as shown in Equation (13, 14).

(6) Determine whether the remaining energy of the cluster head node satisfies the condition to continue to serve as the cluster head, and if not, reselect the cluster head according to Equation (2).

(7) The nodes in the cluster continue to normalize the data listened to during the work and the eigenvalue extraction is performed on the normalized data by using Equation (11).

The flowchart of the R-IEHOBP algorithm proposed in this paper is shown in Figure 3.

5. **Experimental results and analysis.** In this paper, we conducted simulation experiments using MATLAB R2021a, focusing on coal mine gas concentration monitoring as the application scenario. We applied the R-IEHOBP algorithm to process data collected by sensor nodes positioned beneath the mine, capturing parameters such as CO and CH4. The study utilized 674 sets of methane gas and CO emission monitoring data gathered at the working face of a coal mine in Jincheng, Shanxi Province, in 2022. Table 1 presents selected data on methane gas concentration for specific experimental choices.

In this paper, the number of elephant population populations used is 20, the maximum number of iterations is 20, the number of function evaluations is 5, etc., and the number of hidden layer nodes in the neural network is 16. The remaining network simulation parameters are shown in Table 2. Ultimately, the R-IEHOBP algorithm introduced in

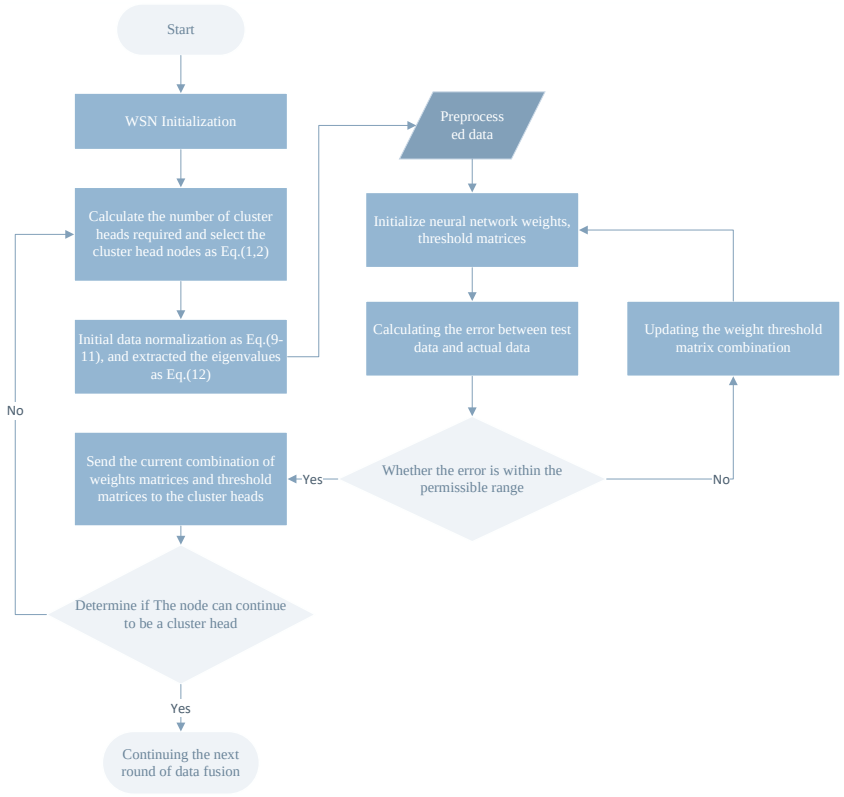

FIGURE 3. RACH structure diagram

TABLE 1. The partial methane concentration data

$CH_4(m^3/t)$	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
x_1	2.19	2.06	2.24	2.31	2.53	2.87	2.82	3.53
x_2	2.23	2.21	2.28	2.35	2.42	2.59	2.97	3.41
x_3	2.17	2.33	2.58	2.68	2.83	3.03	3.58	3.68
x_4	2.21	2.32	2.58	2.68	2.83	3.03	3.58	3.68
x_5	2.23	2.46	2.51	2.45	2.73	2.92	3.25	3.23
x_6	2.18	2.15	2.17	2.24	2.49	2.78	2.87	3.56
x_7	2.12	2.29	2.46	2.55	2.62	2.77	2.86	3.38
x_8	2.23	2.26	2.29	2.37	2.65	2.73	2.69	3.54

this paper is subjected to a comparative analysis with algorithms presented in existing literature (references 22-24). The evaluation is conducted through simulation experiments, focusing on key metrics such as the number of valid packets received, data transmission accuracy, and the average residual energy of the nodes, as reported in references [22, 23, 24].

5.1. **Comparison of valid packets received by aggregation nodes.** In this paper, we evaluate the efficacy of various algorithms in identifying valid packets by comparing the number of valid feature packets transmitted to the base station nodes. This comparison is conducted under an equal number of iteration rounds, considering both the absolute count of valid packets and their growth rate. A higher number of valid data received by the aggregation node within the same number of rounds indicates better algorithm performance.

TABLE 2. The simulation parameter settings for the network

Simulation parameter categories	Simulation Value Setting
Network Emulation Scope $/m^2$	100×100
Number of nodes	200
Node initial energy$/J$	0.5
Transmission consumption$/(nj \cdot bit^{-1})$	45
reception consumption$/(nj \cdot bit^{-1})$	30
Maximum simulation rounds	2000
packet size$/(bits)$	64
simulation time$/s$	500

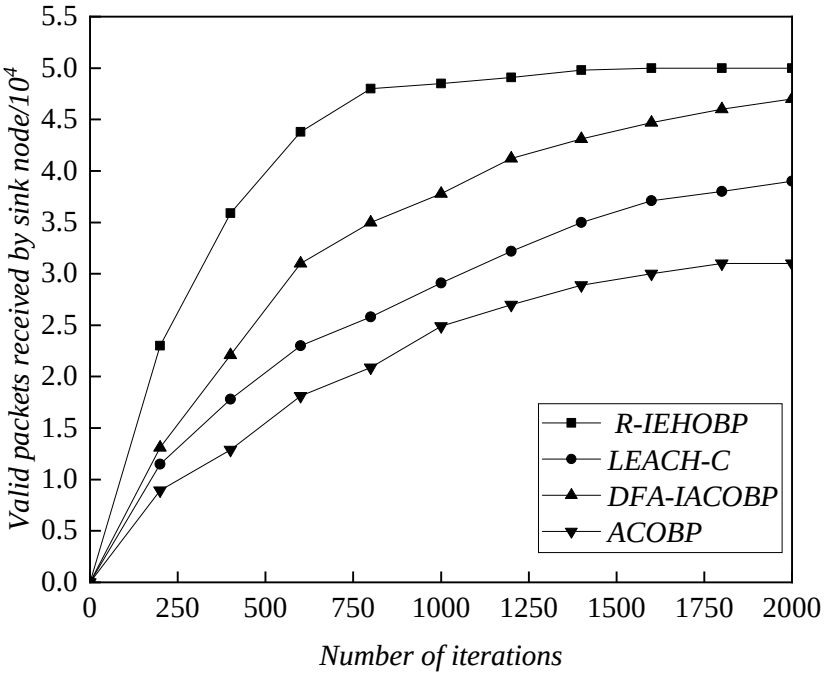

FIGURE 4. Comparison of redundant data rejection capabilities

In Figure 4, the number of featured packets received by the aggregation node is depicted, comparing the usage of the R-IEHOBP algorithm with the other three fusion algorithms. As shown in the figure, the number of valid packets received by the aggregation node using the R-IEHOBP algorithm is more than the aggregation node using the other three algorithms at any given time. The R-IEHOBP algorithm demonstrates notable superiority in packet reception compared to the ACOBP algorithm. Specifically, when the iterations reach 250 times, the R-IEHOBP algorithm receives 2.57 times more valid packets than the ACOBP algorithm. Moreover, with iterations extended up to 1250 times, the R-IEHOBP algorithm achieves a number of valid packets similar to the final count. In comparison, the other three algorithms have to go to 1450, 1600, and 1200 times to achieve the same result. The final results indicate that the algorithm proposed in this paper outperforms

the remaining three algorithms in both the final number of received packets and processing speed.

5.2. **Data Accuracy Comparison.** Ideally, a wireless sensor network should work with 100% data accuracy, but the pursuit of high accuracy means that the rest of the nodes in the network have to bear a higher load. Due to the special characteristics of the working environment in the air-mining zone, the lower the accuracy of the data after the final fusion of the nodes, the higher the possibility of triggering safety problems, so it is necessary to strictly control the accuracy of the data while considering the life cycle of the system in the complex under-mining environment.

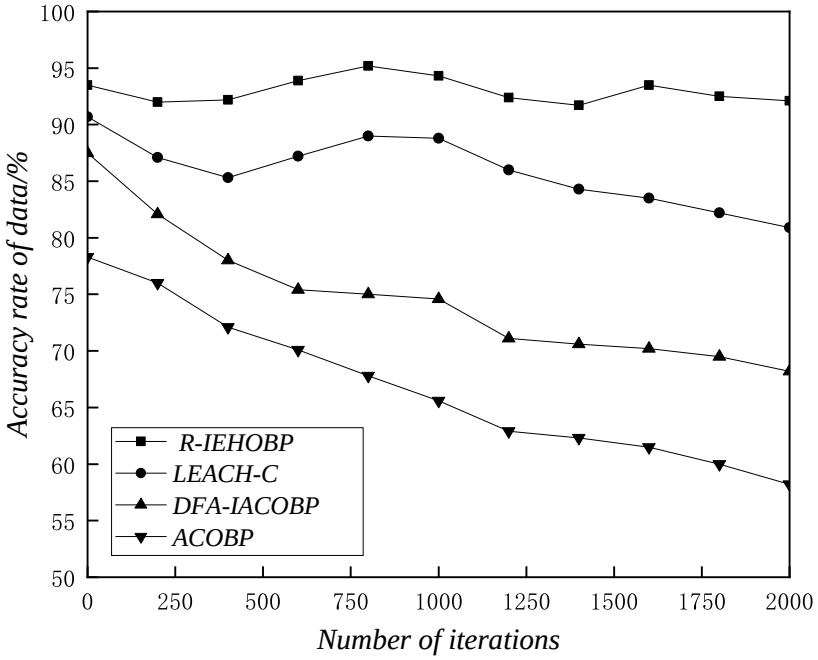

FIGURE 5. Data Accuracy Comparison

As shown in Figure 5, the data accuracy of the R-IEHOBP algorithm in the wireless sensor network simulation experiment is finally maintained at about 95%, while the average accuracy of the LEACH-C algorithm is 88.3%, the average accuracy of the DFA-IACOBP algorithm is 74.5% and the average accuracy of the ACOBP algorithm is 66.7%. The results show that the data accuracy of the R-IEHOBP algorithm is consistently better than the other three algorithms during the simulation experiments.

5.3. **Comparison of average residual energy of nodes.** Wireless sensor nodes carry limited battery energy and are not equipped to recharge or replace batteries. The death of a node creates a blind zone for monitoring, leading to incomplete information collection by the system. The energy hole created by node energy depletion will create an additional workload for other nodes. So the level of node energy consumption is one of the important indicators for evaluating the performance of WSN, the more the average remaining energy of the nodes under the same working time and working environment, the lower the energy consumption of the system under the algorithm and the longer the life cycle of the system.

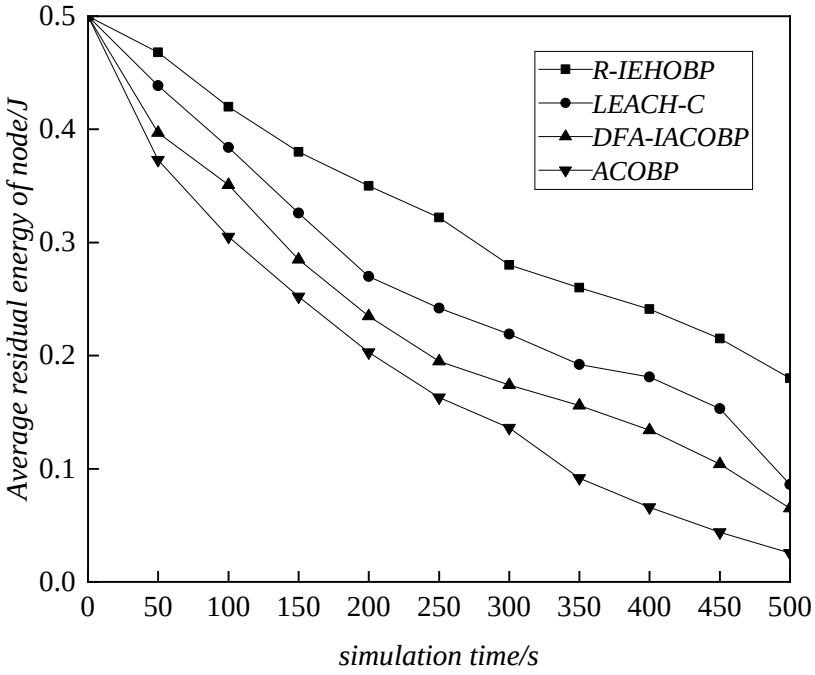

FIGURE 6. Comparison of average residual energy of nodes

As shown in Figure 6 when the simulation time is carried out for 500 seconds, the average residual energy of the nodes of each algorithm is R-IEHOBP algorithm (0.18j), LEACH-C algorithm (0.085j), DFA-IACOBP algorithm (0.065j), and ACOBP algorithm (0.025j). The algorithm proposed in this paper has more residual node energy for the same operating time. And with the simulation experiment, the remaining three algorithms' node residual energy decreases faster, this is because, with the increase of node energy consumption, the WSN needs to replace the cluster head node frequently in order to achieve a more uniform utilization of individual node energy. and the simulation experiments proved that the R-IEHOBP can utilize the nodes of each node more evenly. Simulation experiments demonstrate that R-IEHOBP can utilize the energy of each node in a more balanced way, with lower average energy consumption and longer overall network lifetime than the remaining three algorithms.

6. **Conclusion.** Aiming at the problem that wireless sensor networks have a large amount of redundant data in the working process, which reduces the service life of the network. In this paper, we combine the radial clustering structure of wireless sensor networks with the neural network structure and propose a WSN data fusion algorithm based on radial clustering and swarm neural networks. The algorithm uses the improved EHO algorithm to optimize the initial threshold of the neural network, and then uses the faster convergence speed of the hierarchical neural network to perform the secondary processing of the data, and finally delivers the fused data to the convergence node.

Simulation experiment results show that the R-IEHOBP algorithm proposed in this paper reduces the propagation of redundant data in the network, improves the accuracy of the fused data, and extends the network lifetime. Subsequent tests confirm that the R-IEHOBP algorithm proposed in this paper still maintains good data fusion capability, high

data accuracy and fast data uplink speed when dealing with real-time monitoring systems constructed by multiple monitoring devices. Uncertainty in the selection of thresholds in the two-level neural network leads to fluctuations in the accuracy of the fused data, and the single-quadrant outlier approach still fails to satisfy the demand in the global search, so this is a direction for future research.

Acknowledgment. This work is partially supported by the National Natural Science Foundation of China (No.61872126).

REFERENCES

[1] P. Rawat, K. Singh, H. Chaouchi, J. Bonnin, "Wireless sensor networks: a survey on recent developments and potential synergies," *Journal of Supercomputing*, vol. 68, no. 1, pp. 1-48, 2014.

[2] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, M. Gidlund, "Industrial Internet of Things: Challenges, Opportunities, and Directions," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4724-4734, 2018.

[3] G. Anastasi, M. Conti, M. D. Francesco, "Energy conservation in wireless sensor networks: A survey," *Ad Hoc Networks*, vol. 7, no. 3, pp. 537-568, 2009.

[4] L. Y. Yu, N. Wang, W. Zhang, "Neural network-based data fusion model in wireless sensor networks," *Computer Science*, vol. 35, no. 12, pp. 43-47, 2008.

[5] A. A Nasir, X. Y. Zhou, S. Durrani, et al, "Relaying Protocols for Wireless Energy Harvesting and Information Processing," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3622-3636, 2013.

[6] X. Hou, D. Zhang, M. Zhong, "Research on data fusion algorithm for wireless sensor networks based on event-driven and neural networks," *Journal of Sensing Technology*, vol. 27, no. 1, pp. 142-148, 2014.

[7] J. W. Tian, Y. Liu, W. F. Zheng, "Smog prediction based on the deep belief - BP neural network model (DBN-BP)," *Urban Climate*, vol. 2022, 101078, 2022.

[8] J. Wang, W. Liu, Y. Cheng, "WSN data fusion model based on TEEN protocol and BP neural network," *Computer Application Research*, vol. 34, no. 8, pp. 2486-2489, 2017.

[9] G. Wang, X. Y. Liu, G. G. Wu, Y. Guo, M. Simin, "Research on Data Fusion Method Based on Rough Set Theory and BP Neural Network," *2020 International Conference on Computer Engineering and Application (ICCEA), Guangzhou, China*, 2020, pp. 269-272.

[10] A. Ayhan, U. Y. Huseyin, M. O. Ahmet, T. Bulent, "Neural network based instant parameter prediction for wireless sensor network optimization models," *Wireless Networks*, vol. 25, no. 6, pp. 3405-3418, 2019.

[11] L. Y. Sun, X. X. Huang, W. Cai, "Neural network-based data fusion algorithm for wireless sensor networks," *Journal of Sensing Technology*, vol. 24, no. 1, pp. 122-127, 2011.

[12] H. Gao, L. Gao, C. Zhou, "Research on neural network training algorithm based on particle swarm optimization," *Journal of Electronics*, vol. 25, no. 9, pp. 1572-1574, 2004.

[13] K. Cao, C. Tan, H. Liu, "Data fusion algorithm for wireless sensor networks based on improved gray wolf algorithm optimized BP neural network," *Journal of University of Chinese Academy of Sciences*, vol. 39, no. 2, pp. 232-239, 2022.

[14] X. W. Yu, Q. Liu, X. Y. Li, "BP neural network WSN data fusion algorithm based on improved ant colony," *Journal of Beijing University of Posts and Telecommunications*, vol. 41, no. 4, pp. 91-96, 2018.

[15] K. Koutini, H. Eghbal, M. Dorfer, G. Widmer, "The Receptive Field as a Regularizer in Deep Convolutional Neural Networks for Acoustic Scene Classification," *2019 27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain*, vol. 2019, 10.23919/EUSIPCO.2019.8902732, 2019.

[16] T.-S. Pan, T.-T. Nguyen, T.-K. Dao, and S.-C. Chu, "An Optimal Clustering Formation for Wireless Sensor Network Based on Compact Genetic Algorithm," *2015 Third International Conference on Robot, Vision and Signal Processing (RVSP), Kaohsiung, Taiwan*, vol. 2015, 10.1109/RVSP.2015.77 , 2015.

[17] T. T. Nguyen, J. S. Pan, and T. K. Dao, "A compact bat algorithm for unequal clustering in wireless sensor networks," *Applied Sciences (Switzerland)*, vol. 9, no. 10, p. 1973, 2019

[18] H. J. Wang, T. Jin, H. Wang, "Application of IEHO–BP neural network in forecasting building cooling and heating load," *Energy Reports*, vol. 8, no. 03, pp. 455-465, 2022.

[19] S. Qin, X. Xue, "A Two-Layer Recurrent Neural Network for Nonsmooth Convex Optimization Problems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 6, pp. 1149-1160, 2015.

[20] H. Singh, B. Singh, M. Kaur, "An improved elephant herding optimization for global optimization problems," *Engineering with Computers*, vol. 38, no. 6, pp. 3489-3521, 2021.

[21] T. Eva, C. -H Romana, A. Adis, "Chaotic elephant herding optimization algorithm," *2018 IEEE 16th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2018, pp. 000213-000216.

[22] C. Hang, G. Li, Y. Z. Xie, "WSN data fusion algorithm based on non-uniform clustering and ant colony neural network," *Journal of Sensing Technology*, vol. 33, no. 10, pp. 1483-1488, 2020.

[23] H. Rahman, N. Ahmed, I. Hussain, " Comparison of data aggregation techniques in Internet of Things (IoT)," *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India*, 2016, pp. 1296-1300.

[24] H. Ding, "Data fusion algorithm for WSN networks based on ant colony optimization," *Journal of Shenyang University of Technology*, vol. 42, no. 2, pp. 208-212, 2020.

[25] S. Zhou, S. Kumari, C.-M. Chen, "Transferability of Adversarial Attacks on Tiny Deep Learning Models for IoT Unmanned Aerial Vehicles," *IEEE Internet of Things Journal*, vol.2024, 10.1109/JIOT.2023.3329954, 2024.

[26] T.-Y. Wu, L. Yang, Z.Y Lee, S.-C. Chu, S. Kumari, S. Kumar, "A Provably Secure Three-factor Authentication Protocol for Wireless Sensor Networks," *Wireless Communications and Mobile Computing*, vol. 2021, 5537018, 2021

[27] Y. X. Duan, C. G. Liu, "Elephant herding optimization based on Orthogonal opposition-based learning and Sugeno inertia weights," *2021 International Conference on Artificial Intelligence and Electromechanical Automation (AIEA), Guangzhou, China*, 2021, pp. 263-268.

[28] H. Singh, B. Singh, M. Kaur, "An improved elephant herding optimization for global optimization problems," *Engineering with Computers*, vol. 38, pp. 3489-3521, 2021.

[29] W. Li, G. G. Wang, "Elephant herding optimization using dynamic topology and biogeography-based optimization based on learning for numerical optimization," *Engineering with Computers*, vol. 38, pp. 1585-1613, 2021.