IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 4 October 2023, accepted 8 November 2023, date of publication 14 November 2023,
date of current version 21 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3332914

RESEARCH ARTICLE

Fuzzy Clustering and Routing Protocol With Rules Tuned by Improved Particle Swarm Optimization for Wireless Sensor Networks

LIU YUEBO[1], YU HAITAO[2], LI HONGYAN[3], AND LIU QINGXUE[1]
[1]College of Computer Engineering and Artificial Intelligence, Jilin University of Architecture and Technology, Changchun 130114, China
[2]Department of Information Center, Jilin Communications Polytechnic, Changchun 130015, China
[3]School of Management, Changchun Information Technology College, Changchun 131103, China

Corresponding author: Liu Yuebo (Liuyuebo01@163.com)

This work was supported in part by the Science and Technology Development Project of Jilin Province under Grant 20210203103SF, and in part by the Science and Technology Project of Jilin University of Architecture and Technology under Grant 2019012.

ABSTRACT Fuzzy clustering and routing protocols have been proven to improve energy efficiency, extend network scalability, increase network throughput, balance network load as well as prolong network lifetime. However, rules defined manually according to field experts are impossible or impractical to achieve the optimal solution for a Fuzzy Inference System (FIS). Therefore, a Novel Fuzzy Clustering and Routing Protocol called NFCRP is proposed in this paper by using an improved Particle Swarm Optimization (PSO) algorithm to tune the fuzzy rules. Firstly, one FIS is used to complete clustering based on effective input parameters including residual energy, node degree deviation, and distance to centrality, thereby forming optimal clusters and minimizing the intra-cluster energy consumption. Secondly, the other FIS is adopted to perform routing with descriptors residual energy, distance to BS, and data load deviation, hence addressing the inter-cluster energy consumption. Finally, the rules of both FISs are tuned by an improved PSO algorithm whose parameters are updated by introducing chaotic mapping and adaptive inertia weight. Simulation experiments were conducted to verify the performance of NFCRP against LEACH, EFUCA, EEFUC, FBCR and FMSFLA. According to the results, the average network lifetime of NFCRP increased by 79.59%, 47.99%, 50.35%,15.66 and 13.04%, compared to LEACH, EEFUC, EFUCA, FBCR and FMSFLA. For the average standard deviation of CH's traffic load, NFCRP decreased it by 29.29% over EEFUC, 31.42% over EFUCA, and 25.28% over FMSFLA. For network throughput, NFCRP outperformed LEACH, EEFUC, EFUCA, FBCR and FMSFLA by 16.87%, 46.52%, 48.18%, 29.97 and 71.79%. In addition, NFCRP also reduced energy consumption by 53.95%, 23.76%, 38.72%, 15.71 and 27.18% as compared to LEACH, EEFUC, EFUCA, FBCR and FMSFLA, respectively.

INDEX TERMS Clustering and routing, fuzzy inference systems, particle swarm optimization, energy balance, wireless sensor networks.

I. INTRODUCTION

Wireless sensor network (WSN) is a major technology used in Internet of Things (IoT), which has been used for tremendous applications such as military, industry, agriculture, aerospace, transportation, and so on [1]. Usually, a WSN consists of many tiny inexpensive sensor nodes deployed in hostile or inaccessible environment with non-rechargeable and irreplaceable batteries, whose purpose is to collect information from the environment of interest. Accordingly, energy saving to maximize the network lifetime is highly required for practical applications of WSNs. During the last decades, lots of clustering and routing protocols have been proposed to resolve the problem, which are considered as the most efficient method till now [2]. Clusters are formed to organize the randomly deployed nodes in clustering and routing

The associate editor coordinating the review of this manuscript and approving it for publication was Diego Oliva.

protocols, which contain Cluster Heads (CHs) and Cluster Members (CMs). Generally, CHs are much more important for the network because of their extra capabilities such as data collection, aggregation, forwarding and cluster management than CMs. For a CM, it only transfers the sensed data to its nominated CH in allotted timeslot, or else it enters a sleep state to save energy [3]. Typically, source data is sent to the Base Station (BS) in multi-hop or single hop communication mode. Of course, the multi-hop mode is popular for its advantages such as low transmission range equipment, low contention and attenuation losses, scalability and energy efficiency [4].

Clustering and routing protocols are considered as Non-deterministic Polynomial (NP)-hard optimization problems, whose optimal or near optimal solutions can be determined by using intelligent computing methods such as fuzzy logic [5], genetic algorithm [6], particle swarm optimization [7], or hybridization of fuzzy logic and particle swarm optimization [8], fuzzy logic and grasshopper optimization [9], honey badger and African vulture optimization [10], calf search algorithm and ant colony optimization [11], cuckoo search and rider optimization algorithm [12]. Especially, fuzzy logic can handle the uncertainties inherent in clustering and routing better than its alternatives [4], [5]. Furthermore, fuzzy logic can obtain more flexibility than crisp logic, and achieve the optimal solution by providing a better combination of input parameters as well. Therefore, fuzzy logic has been widely used for CH selection [8], maintenance cycle determination [13], competition radius calculation [14], routing [15], or hybrid [16]. Usually, in fuzzy clustering and routing protocols, FIS is used to form best clusters and find optimal routes simultaneously by integrating different parameters, which includes inputs, inference engine, fuzzy rule, membership function, output, and so on.

In order to make better decision, the parameters in FIS should be well integrated by assigning appropriate membership functions and reasoning the result(s) by setting different fuzzy rules. Residual energy, node degree, distance to centrality are usually used for input parameters to make decision on CH selection, while residual energy, distance to BS, data load are always utilized to find the optimal routes. As for inference engine, Mamdani inference system is the most popular. In addition, trapezoidal and triangle membership functions with different ranges are used to represent the fuzzy inputs and output. Especially, fuzzy rules have a significant impact on decision results. Therefore, traditional manual rule generation based on expertise and practice is almost impossible to make the best decision, and especially not suitable for different applications. Thus several optimization algorithms such as squirrel search algorithm [17], shuffled frog algorithm [4], artificial bee colony [18], and PSO [19] have been adopted to tune the rules. Due to its simplicity, fast convergence and easy implementation, the PSO algorithm can provide a more optimized solution for WSNs than the others [7], [19], [20].

PSO is a commonly used optimization algorithm whose inspiration comes from birds' searching food [21]. In PSO, a population of candidate solutions is represented by a swarm of particles which moves in a search space to solve a problem. The performance of the particle is based on its position. Initially, the position of each particle is randomly generated. In each iteration, a particle updates its position and velocity according to its local best and global best of all the particles so as to reach the global optimum. Moreover, a fitness function is defined to evaluate each particle considering different parameters. In WSNs, PSO has been used for clustering [22], routing [23], clustering and routing [24], target localization and tracking [25], sleep scheduling [26], intrusion detection [27], congestion control [28], and so on [20], [29], [30]. In existing clustering and routing protocols based on PSO, it is used to reach two objectives, i.e. selecting best CHs and finding optimal routing paths. In the end, the total energy consumption is reduced and the network lifetime is extended. PSO has many advantages over other alternatives optimization techniques to solve NP-hard problems [7]. In addition, Compared with the other optimization approaches, PSO has shown greater excellence in exploration and exploitation applications [19]. Moreover, PSO has also been widely used to solve the optimization problems for membership function adjustment and fuzzy rules tuning for FISs, with many advantages such as easy implementation, availability to escape from local optima, and quick convergence [30], [31].

In this paper, a novel fuzzy clustering and routing protocol with rules tuned by an improved particle swarm optimization algorithm called NFCRP is proposed to enhance the energy efficiency, extend the network lifetime as well as resolve the hot-spot issue. In NFCRP, clusters are formed by using one FIS with descriptors residual energy, node degree and distance to centrality. Moreover, the other FIS uses parameters such as residual energy, distance to BS, and data load to find the optimal routes for each CH. Remarkably, the rules of both FISs are generated by an improved PSO algorithm with adaptive inertial weight adjustment. The major contributions of the proposed protocol NFCRP are listed as below.

1. Optimal clusters and routes are generated by using FISs integrated with different parameters to improve energy efficiency, balance energy consumption and mitigate hot-spot issue.
2. Optimal fuzzy rules are determined by using an improved PSO algorithm with chaotic mapping and adaptive inertial weight adjustment so as to maintain the population diversity and balance the capabilities of exploration and exploitation at the same time.
3. Simulations are conducted to demonstrate the superiority of NFCRP over some existing up-to-date protocols in different scenarios and performance metrics, simultaneously.

The rest of this paper is organized as follows: Literatures about fuzzy clustering and routing protocols are briefly reviewed in Section II. The system model including network model and energy model is addressed in Section III. Section IV introduces the proposed NFCRP protocol in detail. In Section V, simulation analysis and comparison with the existing protocols are presented to evaluate the proposed protocol. Finally, conclusion is stated and future directions are indicated in Section VI.

II. LITERATURE REVIEW

Clustering and routing protocols have demonstated strong capabilities in energy efficiency, scalability, throughput, and other aspects since the first proposing of LEACH [3]. In LEACH, two phases, i.e. setup phase and steady phase, are used to form clusters and transfer data, respectively. In setup phase, each node makes decision to be CH based on probabilistic model, in other words, the sensor node becomes a CH if its randomly assigned number between 0 and 1 is lower than the threshold value. Otherwise, it becomes a normal node. Once the CHs are determined, messages are broadcast to announce their status. At the same time, each normal node joins its corresponding CH according to the strength of the received signal from CHs. Moreover, a request message is also sent to its CH telling its CM status. After establishing the clusters, every CH sends to its members a TDMA (Time Division Multiple Access) schedule for data transmission. In steady phase, each CM sends its sensed data to its CH using the nominated timeslot in the TDMA schedule. Afterwards, the CH fuses the data from its CMs, and forwards these data to the BS directly by using CDMA (Code Division Multiple Access) code to avoid the collision. In LEACH, the number of packets is decreased by forming clusters and selects CHs. In addition, interference between nodes are also avoided by using TDMA and CDMA mechanism, which reduces the network energy consumption as well. However, its disadvantages deteriate the network performance, including random selection of CHs without considering parameters such as residual energy and node degree, single-hop data forwarding between CH and BS, and CM joining the nearest CH to form cluster based on received signal strength. Therefore, a good many of improved clustering and routing protocols have been proposed to enhance the overall performance of the network [32], [33]. However, as a well known NP-hard problem, traditional clustering and routing protocol is inefficient or even impossible to solve this non-polynomial issue, not to mention the uncertain, dynamic and self-organizing nature of the network [2], [6]. Exactly, fuzzy logic exploits the uncertainty and dynamics associated with the factors that affect the network lifetime of WSNs and improves their performance in real-world applications [34], [35]. Thus, only fuzzy based clustering and routing protocols related to the proposed NFCRP are put into consideration here.

In LEACH-ERE [36], a probabilistic based mechanism like LEACH is used to select the candidate CHs at first, then the candidates use fuzzy logic to calculate the chance of being a CH considering two inputs residual energy and expected residual energy. The LEACH-ERE protocol can effectively improve the network lifetime. However, it may result in non-uniformly distributed clusters due to its probabilistic selection of the candidate CHs and neglect of other parameters. Accordingly, in FBCR [37], three parameters including residual energy of the node, node distance to the center of region, and node angle to the BS are used as the input of the FIS which can maintain energy balance and better CHs selection and rotation. Besides, the tasks such as selecting CHs in regions, determing the region of each node, calculating the distance of each node to the center of its region, and the angle of each to the BS are all competed by the BS so as to reduce the burden of choosing CHs for nodes. Simulation results show that FBCR can effectively improve the energy efficiency. However, similar to LEACH-ERE, the CHs bearing more workload easily lead to premature death. So in [38], K-means and PSO are used to form appropriate clusters firstly, and then both fuzzy logic systems with descriptors energy level, distance to cluster's center, distance to BS and energy level, distance to PCH respectively are utilized to select primary CH (PCH) and secondary CH (SCH) for tasks sharing to enhance network lifespan and throughput. However, the hot-spot issue is still not considered. Therefore, in MOFCA [39], a multi-objective fuzzy clustering algorithm is proposed to solve these problems. The CHs are determined by energy-based fuzzy competition among the candidate CHs initially selected by using a probabilistic model similar to LEACH-ERE. Three inputs are used for fuzzy inference system in MOFCA, namely residual energy, distance to sink, and node density, so as to estimate the competition radius and select the final CHs. However, node's centrality is not considered in MOFCA, which may increase the intra-cluster energy consumption. Accordingly, In OFCA [8], the fuzzy inference system is used to calculate the confidence factor for each node which indicates the goodness of its becoming a CH. Three parameters are considered as inputs of the fuzzy inference systems, namely residual energy, distance to BS and concentration. The node with maximum confidence factor is selected as CH. Furthermore, the communication radius of CHs are determined by the same fuzzy inference system based on distance to BS. Similar to MOFCA, the CHs closer to BS are set with smaller communication radius than that of the farther ones, because the CHs closer to BS are overloaded with inter-cluster traffic. In addition, an energy efficient routing path is established for multi-hop data forwarding by using PSO with a novel fitness function. Simulation results prove that OFCA attains longer network lifetime and transfers more packets to BS. However, quality of transmission link and the responsibility of node in preceding rounds are not considered in OFCA, which undoubtedly affects its practical application and reduces the overall performance of the network. Thus, In FLEAC [9], fuzzy inference system with five descriptors residual energy, average intra-cluster distance, compaction degree, packet drop probability and node history is utilized to select appropriate CHs. Then, the node having the best output is selected as a CH. Moreover, improved grasshopper optimization algorithm is employed to select the optimal relay nodes for data forwarding, in order to alleviate excessive energy consumption of CHs. FLEAC outperforms other comparative clustering protocols, and it minimizes and balances the energy consumption among the nodes, and the network lifetime is largely extended. However, the fuzzy rules are fixed, which are not suitable for different applications. Hence, in DPFCP [30], a fuzzy logic system with rules tuned by PSO is presented to select the optimal CHs. Its input parameters include residual energy, node degree, distance to BS, and distance to the centrality. Moreover, on demand re-clustering mechanism is adopted to maintain clusters so as to further save energy consumption. Simulation results show that DFPCP can efficiently balance energy consumption and extend network lifetime.

In MLSEEP [40], fuzzy logic system is utilized to select the best next-hop CH for each CH from its neighbor CHs so as to extend the network lifetime and decrease the network overload. The following three descriptors queue length, distance to the BS and residual energy are used for inputs of the fuzzy logic system. Then, the CHs closer to the BS with higher residual energy and queue length has bigger probability to be selected for next-hop CH. However, probability based CH selection is apt to nominate inappropriate CHs to form clusters in MLSEEP. Accordingly, unbalanced energy consumption occurs in clusters and the network lifetime is decreased in the end. So, in E-FEERP [41], the PSO algorithm instead of probability based method is used to form clusters considering distance to neighbors, average energy of neighbors and number of neighbor clusters. Once the clusters are formed, a fuzzy logic system with descriptors battery power, average node density, average distance, communication quality is adopted to determine the optimal next hop of CHs. E-FEERP has been demonstrated to improve the network lifetime, throughput, and packet delivery ratio. However, the hot spot issue is not considered in E-FEERP. Accordingly, in DAPFL [15], a fuzzy logic system is used to find the optimal next-hop CH for each CH, whose inputs includes residual energy, distance to BS and data length. The neighbor CH with the best chance value is selected as the final next-hop CH. Moreover, DAPFL solves the hot-spot issue by adjusting the energy consumption among the CHs according to their data load. In addition, affinity propagation rather than probability based mechanism in MLSEEP is used to determine the number of clusters and select the appropriate CHs simultaneously based on residual energy, distance between nodes. Simulation results reveal the effectiveness of DAPFL with respect to network energy consumption, standard deviation of residual energy, network throughput and lifetime. But centralized decision on cluster formation may reduce the overall performance of the network. Consequently, in EFUCA [14], fuzzy logic is used to take intelligent decision for each node to become a CH. Besides the remnant power, nearness to BS, average distance to communicating nodes are considered as inputs to the fuzzy logic system for calculating rank and competition radius in the clustering stage. To further prolong the network lifetime, another fuzzy logic system is utilized to complete the next-hop choice for efficient data forwarding, whose inputs are next-hop rank, nearness to next-hop and distance reduced to BS. The experimental results prove the improved performance of EFUCA with its comparatives in extended lifetime, protracted stability period, and decreased average energy consumption. However, a normal node joins a cluster only considering its distance to CH, resulting in decreased network lifetime due to the unbalanced intra-cluster energy consumption. Hence, in EEFUC [16], different fuzzy logic systems are used to make decisions on CH selection, competition radius calculation, joining the appropriate CH, and finding next-hop CH based on various input parameters. Residual energy, node density and distance to BS are taken as fuzzy inputs to calculate the competition radius. For CH selection, residual energy of a node is the decisive factor. When the residual energy of nodes is the same, the second fuzzy logic system is employed to select the best one as the final CH with descriptors node density and distance to BS. Moreover, the third fuzzy logic system with inputs residual energy and distance to CH is adopted to make a node join an appropriate CH, so as to form optimal clusters. Once the clusters are established, the forth fuzzy logic system is used to find the proper next-hop CH for each CH, whose inputs include residual energy of neighbor CH, distance to neighbor CH and DOP which is the distance to the line from CH to BS. Simulation results in various scenarios show that EEFUC can not only extend the network lifetime largely but also effectively reduce the network energy consumption. However, the fuzzy rules are defined manually, resulting in fixed and nonadjustable performance.

In SIF [42], fuzzy c-means and fuzzy logic system are utilized to form balanced clusters and select appropriate CHs, respectively. Moreover, a hybrid swarm intelligence algorithm based on firefly algorithm (FA) and simulated annealing (SA) is utilized to optimize the fuzzy rules of SIF. Firefly algorithm is employed to obtain the optimal fuzzy output for each rule in the search space with its global exploration ability. And then, the final best solution of firefly algorithm is used for initial solution of simulated annealing. Iteratively, based on its local search ability, simulated annealing can achieve the optimal solution in the end. In addition, the fitness function is defined to evaluate the solution based on network lifetime FND (First Node Dies), HND (Half Nodes Die) and LND (Last Node Dies). Specifically, the FA-SA algorithm is performed to tune SIF once before the network operation. Therefore, the optimization procedure does not boost any computational complexity and delay in the data transmission phase. Experimental results over different networks show that SIF is superior to the comparative protocols in terms of forming balanced clusters and extending the network lifetime. Similar to SIF, LEACH-SF [18] also uses fuzzy c-means and fuzzy logic system to form balanced clusters and select the appropriate CHs. And the difference is that the artificial bee colony (ABC) algorithm is presented to adaptively tune the fuzzy rules in LEACH-SF so as to prolong the network lifetime for each application. The feasible solutions in ABC algorithm are represented as strings of length as same to the number of controllable parameters. A fitness function is constructed to evaluate the quality of each bee based on the network lifetime FND, HND and LND. Through population updating based on the randomly generated initial population, the best solution will be reached in the end. Same as SIF, the iteratively-based optimization algorithm is performed to tune fuzzy rules only once before the network operation in LEACH-SF. Simulation results validate that LEACH-SF can minimize the intra-cluster distances, maximize the network lifetime and throughput at the same time. In SF-MPSO [19], the differences from SIF and LEACH-SF are the used optimization method and the defined fitness function. PSO with excellent exploration and exploitation ability is utilized to tune the controllable parameters. Each solution represented by a particle undergoes the evaluation process by optimizing the fitness value. The fitness value of each particle is calculated by the designed fitness function considering minimizing the intra-cluster distance, maximizing the network residual energy and the correlation coefficient. Iteratively, the global optimal solution with the greatest fitness value is reached by updating the position and velocity of each particle. It is observed that SF-MPSO performs outstandingly under all simulation scenarios and has longer network lifetime compared with its comparatives. However, fuzzy rules are tuned only for clustering in above mentioned protocols. Therefore, in FMSFLA [4], CHs are selected from the nodes with higher residual energy rates, shorter distance to BS, and shorter distances to neighbors by using a fuzzy logic system. Similarly, the relay nodes are also selected from the CHs with respect to higher residual energy rates, shorter real distances, and lower path loads by employing another fuzzy logic system. Moreover, the rules in both fuzzy logic systems are tuned by using shuffled frog leaping algorithm (SFLA) with a novel fitness function considering network lifetime, average number of packets received by the BS, and average delay of the number of hops. Therefore, the fuzzy rules can be adjusted to meet the application features. Once the CHs and relay nodes are determined, the non-CH nodes join respective cluster with the nearest CH to form clusters. And then non-CH nodes send their data to the corresponding CH which aggregates the data and forwards it to the BS directly or through their relay nodes. Simulation results show that FMSFLA can achieve steady network workload, decrease the network energy consumption, and extend the network lifetime. Comparison of different protocols related to the proposed NPSOP is shown in Table 1.

III. SYSTEM MODEL

A. NETWORK MODEL

In NFCRP, the network consists of n sensor nodes with limited power, storage, computational and communicational capabilities, which are homogeneous. Moreover, the nodes are randomly deployed over the target field with unique ID numbers. Additionally, the nodes and the BS are immobile once being deployed, and the BS has no constraints in terms of power, processing, storage and other resources. The distance among nodes and BS can be obtained by received signal strength, and the transmission power of nodes can be adaptively determined to communicate with each other. For simplicity, only symmetric links are considered for communications among nodes and BS.

All the nodes are formed into clusters, and a CH is selected to manage each cluster, while the BS is responsible for management of the network at the same time. Besides, the BS can locate inside or outside of the field of interest. Like LEACH, round is used for periodic data collection, in which all source nodes send their sensed data to the BS. Round by round, a certain node will die when its energy is fully drained. Similarly, the network dies when all the nodes are lifeless. Therefore, the network lifetime in this paper is also measured by FND, HND, and LND as traditional approaches [4], [18]. The network model can be shown in Fig. 1.

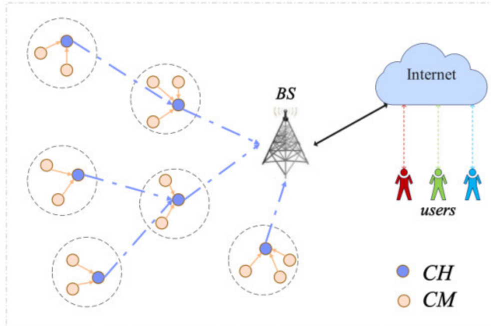


FIGURE 1. The network model of the proposed NFCRP protocol.

B. ENERGY MODEL

Energy of the nodes is usually consumed for data sensing, transmitting, receiving, fusing, and processing. However, the energy drained in sensing and processing can be negligible [1], [43]. Then the energy dissipated for data transmitting from nodes i to j with k bits over distance d can be given in Eq. (1) [30], [44].

$$E_{Tij(k,d)} = \begin{cases} k\times E_{elec}+k\times\varepsilon_{fs}\times d^2, & d < d_0 \\ k\times E_{elec}+k\times\varepsilon_{mp}\times d^4, & d \geq d_0 \end{cases} \quad (1)$$

where d is the Euclidean distance between nodes i and j. d_0 denotes the threshold used to determine either free space (ε_{fs}) or multi-path (ε_{mp}) model adopted, which can be estimated by $d_0 = \sqrt{\varepsilon_{fs}/\varepsilon_{mp}}$. E_{elec} means the electronics energy consumption required to transmit or receive 1-bit. ε_{fs} and ε_{mp} are the amplifier coefficients used for free space and multi-path model, respectively. Similarly, the energy consumption of node i for receiving k-bit data from node j is given in Eq. (2) [30], [44].

$$E_{Rij}(k) = k\times E_{elec} \quad (2)$$

Additionally, the energy dissipated for fusing k-bit data can be calculated according to Eq. (3) [15], [30].

$$E_{DA} = k\times E_{pDb} \quad (3)$$

where E_{pDb} is the energy consumed for fusing 1-bit data.

TABLE 1. Comparison of related protocols.

Protocol	Method	Data forwarding	Node type	Fuzzy inference system				Objectives
				Subject	Input parameters	Rules	Output	
[36]	Distributed	Single-hop	Homogenous	Clustering	Residual energy, expected residual energy	Defined mannually	Chance	Prolong the lifetime
[37]	Centralized	Single-hop	Homogenous	Clustering	residual energy of the node, node distance to the center of region, and node angle to the BS	Defined mannually	Chance	Reducing energy consumption
[38]	Hybrid	Single-hop	Homogenous	Clustering	FIS1:energy level, distance to cluster's center, distance to BS FIS2:energy level, distance to PCH	Defined mannually	PCH selection chance, SCH selection chance	Enhance lifespan and throughput
[39]	Distributed	Multi-hop	Homogenous	Clustering	Remaining energy, distance to the BS, density of the tentative CH	Defined mannually	Competition radius	Energy efficiency, hot spot problem
[8]	Centralized	Multi-hop	Hybrid	Clustering	Residual energy, distance from sink, concentration	Defined mannually	Confidence factor, communication radii	Energy efficiency, hot spot problem
[9]	Centralized	Multi-hop	Homogenous	Clustering	Residual energy, average intra-cluster distance, compaction degree, packet drop probability, node history	Defined mannually	Chance	Energy efficiency
[30]	Distributed	Single-hop	Homogenous	Clustering	Residual energy, node degree, distance to BS, distance to the centroid	Tuned by PSO	Chance	Maximizing network lifetime, balance energy consumption
[40]	Distributed	Multi-hop	Homogenous	Routing	Queue length, distance from the BS, remaining energy	Defined mannually	Probability	Energy efficiency
[41]	Hybrid	Multi-hop	Homogenous	Routing	Battery power, average node density, average distance, communication quality	Defined mannually	Likelihood	Energy efficiency
[15]	Hybrid	Multi-hop	Homogenous	Routing	Residual energy, data length, distance to BS	Defined mannually	Chance	Energy efficiency, energy balance, hot spot problem
[14]	Distributed	Multi-hop	Homogenous	Hybrid	FIS1: Residual energy, closeness to BS, average distance FIS2: Next hop rank, nearness to next-hop, distance reduced to BS	Defined mannually	FIS1: Rank, competition radius FIS2: Cost	Energy efficiency, hot spot problem
[16]	Distributed	Multi-hop	Homogenous	Hybrid	FIS1: Distance to BS, node density FIS2: Distance to BS, remaining energy, node density FIS3: Distance to CH, remaining energy FIS4: Remaining energy, delay distance, distance to optimal point	Defined mannually	FIS1: CH_Chance FIS2: Competition radius FIS3: CM_Chance FIS4: nextCH_Chance	Energy balance, hot spot problem
[42]	Centralized	Single-hop	Homogenous	Clustering	Residual energy, distance to sink, distance to cluster center	Tuned by FA and SA	Z(n)	Prolong the network lifetime
[18]	Centralized	Single-hop	Homogenous	Clustering	Residual energy, distance from sink, distance from cluster centroid	Tuned by ABC	Z(n)	Prolong the network lifetime

TABLE 1. *(Continued.)* **Comparison of related protocols.**

[19]	Centralized	Single-hop	Homogenous	Clustering	Residual energy, distance from cluster centroid, distance from the BS	Tuned by PSO	Z(n)	Maximize the network lifespan
[4]	Centralized	Multi-hop	Homogenous	Hybrid	FIS1: Residual energy, distance from the BS, number of neighboring nodes FIS2: Energy level, real distance between candidate CHs and the BS, the load mean of the route between the candidate CH and the BS	Tuned by SFLA	FIS1:Out_crisp$_{CH}$(n) FIS2: Out_crisp$_P$(n)	Maximize the network lifetime and the number of received packets

Accordingly, the total energy exhausted of a cluster with CH i and m CMs in a round can be given as follows:

$$E_{CHi} = \sum_{j=1}^{m} E_{Rij}(k) + \sum_{j=1}^{m} E_{Tij}(k, d_{ij}) + E_{DA} \tag{4}$$

IV. THE PROPOSED NFCRP PROTOCOL

The NFCRP is a distributed energy balanced and efficient clustering and routing protocol used to maximize the network lifetime and mitigate the hot-spot issue. To this end, fuzzy clustering and fuzzy routing with rules generated by improved PSO are utilized to select the best CHs and find the optimal routes, respectively. Especially, the rules only need to be tuned at the first time performing the fuzzy computation for each CM or CH in a specific application. The framework of FISs in NFCRP is given in Fig.2.

As seen from Fig.2, fuzzy logic is an efficient human decision-making behaviour modelling tool, and relatinal expressions or set of linguistic variables can be used to express the input-outp relations. Usually, FIS comprises of fuzzifier, fuzzy inference engine, a rule base and defuzzifier. The inference engine forms inferences that help to draw inference from the fuzzy rules. The defuzzification unit receives the output and maps fuzzy actions space into a crisp actions space [4], [9]. The detail introduction of NFCRP is described as follows, including fuzzy clustering, fuzzy routing, fuzzy rules optimization, and time complexity analysis.

A. FUZZY CLUSTERING

Here, the FIS is employed to select candidate CHs. Moreover, only the nodes with residual energy higher than the mean energy of the network perform fuzzy computations so as to reduce the computational complexity. Three parameters residual energy, node degree deviation, and distance to centrality are regarded as fuzzy inputs. In order to put the data of different input parameters from different domains in the same domain, the normalization is adopted to map the crisp input data onto another value ranging in [0, 1], which can be expressed as follows:

$$v'_{in} = \frac{v_{in} - v_{in-min}}{v_{in-max} - v_{in-min}} \tag{5}$$

where v_{in-max} and v_{in-min} are the maximum and minimum of input data v_{in}. Generally, the input parameters have a key role in a FIS, which are described as follows:

-$E_{res}(i)$: indicates the residual energy of node i, and the fuzzy linguistic variables for $E_{res}(i)$ are "very less", "less", "medium", "much", and "very much". Without loss of generality, the nodes with more residual energy have a greater chance of being selected as candidate CHs.

-$V_{nd}(i)$: indicates the deviation of node i's node degree to the mean number of neighbors of all the nodes, and its fuzzy linguistic variables are "small", "medium", and "big". Obviously, the smaller the deviation, the greater of the chance of a node to being selected as candidate CH. More importantly, this parameter is used to form uniform clusters.

-$D_{toc}(i)$: indicates the distance of node i to the centrality of its neighbors, and "very near", "near", "normal", "far", and "very far" are considered as its fuzzy linguistic variables. Moreover, the nodes with smaller distance have a better chance to become candidate CHs.

Similar to the traditional approaches [14], [45], trapezoidal and triangular member functions are used for boundary and intermediate fuzzy linguistic variables, respectively, because of their capabilities of providing faster calculation and simpler implementation. After the process of fuzzification of the crisp input values by using the membership of corresponding fuzzy linguistic variables, the variables are fed to the fuzzy clustering inference system which is a Mamdani inference model with simplicity and characteristics. Then, the IF-THEN rules are utilized to calculate the fuzzy output "chance" with fuzzy linguistic variables "very low", "low", "medium", "high", and "very high". Usually, the rules are manually adjusted according to the experts' knowledge and empirical data of a specific application, which is impossible to make optimal decisions on different applications. Therefore, the rules are adaptively tuned by an improved PSO algorithm in NFCRP which will be discussed in subsection C. Based on the generated rules optimized by PSO, the crisp output chance′ of node i is determined by performing defuzzification with Center Of Area (COA) method. The corresponding membership functions for input and output variables are depicted in Fig.3.

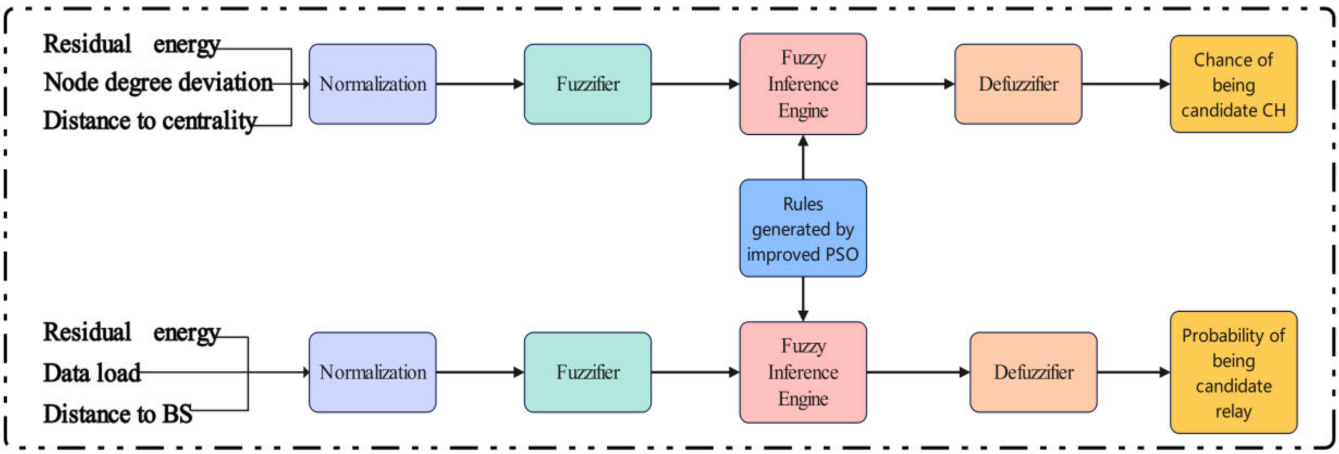


FIGURE 2. **Framework of the FISs in the proposed NFCRP protocol.**

After the calculation of the crisp output *chance′* of each node, a message containing node ID and value of *chance′* is broadcasted by every node. Then the node with higher chance′ and appropriate distance with neighbor CHs is determined as final CH. The possible distance between two adjacent CHs for proper communication is ranging in [0, R], where R is the communication radius of the nodes. Moreover, the final selected CHs announce their CH status by broadcasting messages, and each non-CH node joins its nearest CH as a member while receiving the messages. Otherwise, the non-CH node will make itself as CH in the worst case. Additionally, TDMA is adopted to prevent loss of data in a collision among the clusters. Thus, the energy efficient and uniformly distributed clusters are formed. The processes for fuzzy clustering are described in Algorithm 1.

Algorithm 1 Pseudo Code for Fuzzy Clustering

- **Input**
- $\mathbf{S=\{S_1, S_2, \ldots, S_n\}}$
- number of CHs: m
- **Output**
- An assignment S–>CH
- **Begin**

Calculate E_{res}, V_{nd}, D_{toc}
Step 1:
Normalize E_{res}, V_{nd}, D_{toc} between 0 to 1 using Eq. (5)
Step 2:
Fuzzify by using correlate membership functions
Step 3:
Fuzzy inference by using rules generated by Algorithm 3
Step 4:
Defuzzify the fuzzy values generated by the FIS using COA
$m = \max(\text{chance}')$
Declare the m_{th} sensor node as a CH
Assign the sensor nodes from each cluster to the respective CH

- **End**

B. FUZZY ROUTING

Once the clusters are formed, the member nodes send their sensed data to their respective CH in allotted timeslot, and the CHs aggregate the received data before transferring it to the BS. Usually, single-hop data forwarding of CHs depletes much more energy than multi-hop. In addition, unbalanced energy consumption and hot-spot issue exist in many traditional routing approaches due to the determination of improper relay CHs. In NFCRP, the other FIS is used to make decision on optimal routes finding.

The objectives of fuzzy routing are improving energy efficiency, balancing energy consumption and mitigating hot-spot issue, simultaneously. For any CH, its energy consumption mainly comes from completing three tasks: one is to transfer the fused data within the cluster, the other is to receive data from its adjacent CHs as a relay, and the third is to forward these data, which can be elaborated in Fig. 1. Hence, residual energy, distance to BS and data load deviation should be considered carefully to achieve the above mentioned objectives. Similar to fuzzy clustering, the input parameters are normalized to range in [0, 1] at first. And the fuzzy linguistic variables applied to residual energy E_{res} are also ''very less'', ''less'', ''medium'', ''much'', and ''very much''. For distance to BS D_{toB}, the applied linguistic variables are ''very near'', ''near'', ''normal'', ''far'', and ''very far''. Besides, the input parameter data load deviation V_{DL} uses ''small'', ''medium'', and ''big'' as its fuzzy linguistic variables. Accordingly, the fuzzy output prob has fuzzy linguistic variables ''very low'', ''low'', ''medium'', ''high'', and ''very high''. Also, the boundary and intermediate fuzzy linguistic variables are also use trapezoidal and triangular member functions, respectively. The membership functions used for input and output variables for fuzzy routing are given in Fig. 4.

The IF-THEN rules are also tuned by the improved PSO algorithm discussed in subsection C, and each CH calculates the eligibility of its adjacent CHs in the direction of BS based on the optimized rules. Certainly, the crisp output prob′ can be obtained by using defuzzification of COA method. Thus, the candidate relay with the maximum prob′ is selected as the final relay CH, and the CH forwards its data to the selected relay CH, till to the BS in the end. The processes for fuzzy routing are described in Algorithm 2.

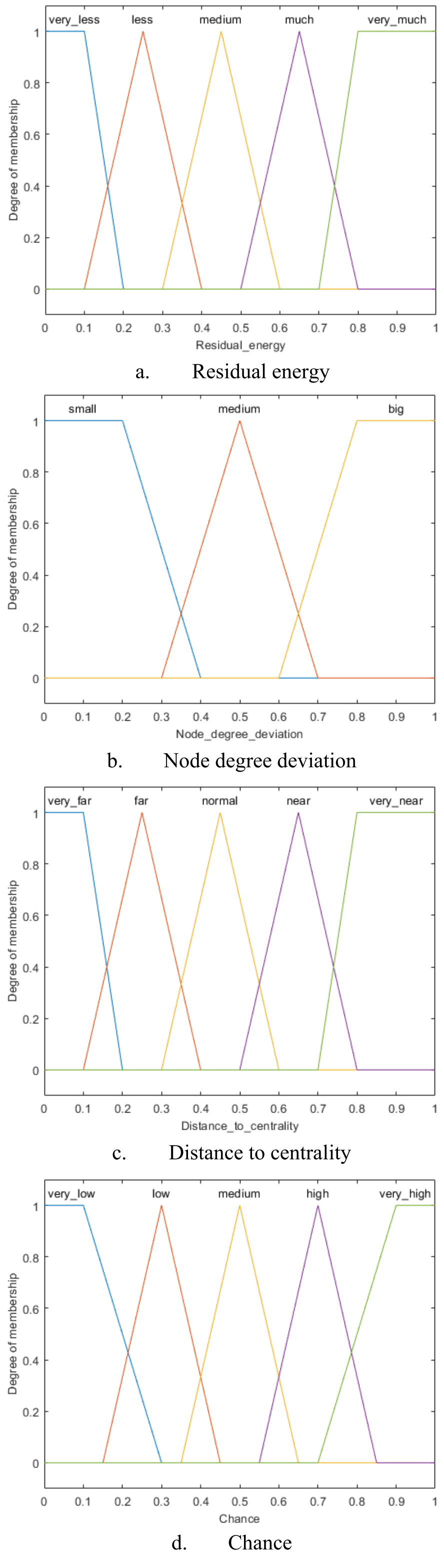


FIGURE 3. Membership function for input and output variables in fuzzy clustering.

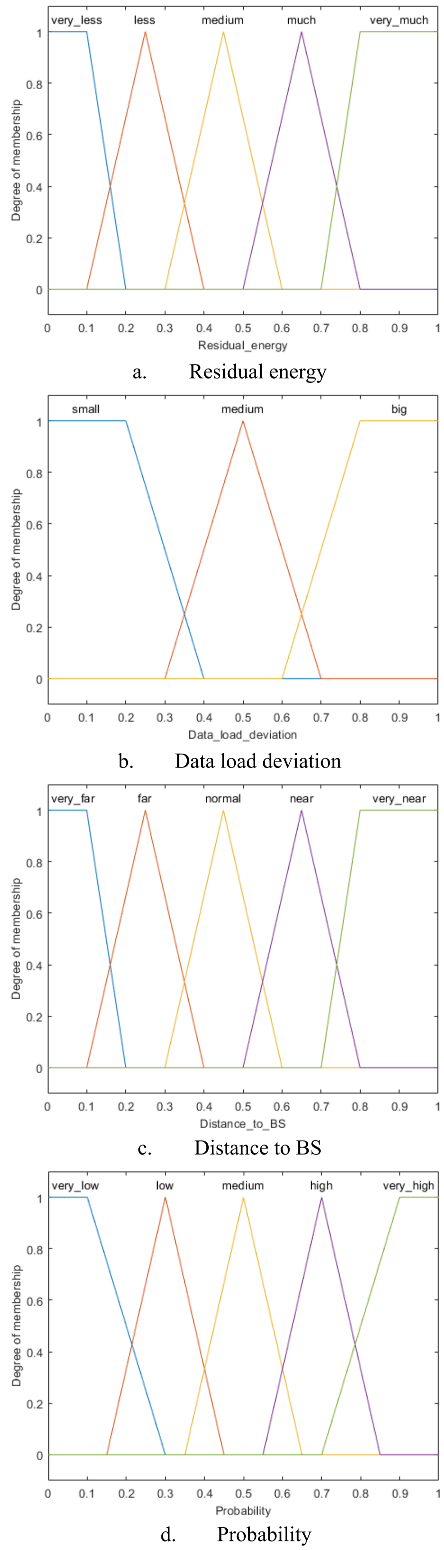


FIGURE 4. Membership function for input and output variables in fuzzy routing.

C. FUZZY RULES OPTIMIZATION

According to the inputs and outputs of the FISs for clustering and routing, their rules can be listed as Table 2.

In the table, the preceding variables in the first line are for fuzzy clustering, and the following ones are for fuzzy routing. Different from traditional fuzzy based approaches,

Algorithm 2 Pseudo Code for Fuzzy Routing

- **Input**
- **H={CH_1, CH_2, ..., CH_m}**
- **Output**
- An assignment H–>next_hopCH
- **Begin**

1: Calculate E_{res}, V_{DL}, D_{toB}
2: **Step 1:**
3: Normalize E_{res}, V_{DL}, D_{toB} between 0 to 1 using Eq. (5)
4: **Step 2:**
5: Fuzzify by using correlate membership functions
6: **Step 3:**
7: Fuzzy inference by using rules generated by Algorithm 3
8: **Step 4:**
9: Defuzzify the fuzzy values generated by the FIS using COA
10: next_hopCH=max(prob′)
11: Assign next_hopCH to the respective CH

- **End**

TABLE 2. IF-THEN rules.

No.	E_{res}/E_{res}	V_{nd}/V_{DL}	D_{toc}/D_{toB}	Chance/prob
1	very less	small	very near	o_1
2	less	small	very near	o_2
3	medium	small	very near	o_3
4	much	small	very near	o_4
5	very much	small	very near	o_5
6	very less	medium	very near	o_6
7	less	medium	very near	o_7
...	...	...	...	...
75	very much	big	very far	o_{75}

the outputs chance/prob o_1-o_{75} are generated by using PSO in NFCRP. Each particle $\boldsymbol{p_i} = \{\boldsymbol{x_{i1}}, \boldsymbol{x_{i2}}, \ldots, \boldsymbol{x_{in}}\}$ indicates a possible solution for the outputs. Furthermore, the chaos theory and adaptive inertial weight adjustment are used to guarantee the population diversity and keep the balance between exploration and exploitation, so as to avoid falling into local optima as much as possible.

Random initialization is adopted to generate the initial population, and the particles' position x_{ik}^t and velocity v_{ik}^t are updated as Eqs. (6) and (7) [21], [23].

$$x_{ik}^t = x_{ik}^{t-1} + v_{ik}^t \tag{6}$$

$$\begin{aligned} v_{ik}^t = \omega \times v_{ik}^{t-1} + c_1 \times r_1(pbest_i - x_{ik}^{t-1}) \\ + c_2 \times r_2 \times (gbest - x_{ik}^{t-1}) \end{aligned} \tag{7}$$

where n denotes the number of particles, and k indicates the dimension of a particle which equals the number of rules. t means the number of iteration, r_1, r_2 are uniformly distributed random numbers ranged in [0, 1]. c_1 and c_2 called learning factors, and ω denotes the inertial weight. $pbest_i$ and gbest represent the best position visited by particle i and overall particles.

In order to increase the probability of finding the optimal solution, the learning factors and inertial weight are optimized by chaos theory and adaptive adjustment, respectively. Firstly, the chaos search is used to obtain the optimal values of c_1 and c_2 by introducing a logistic chaotic function given in Eq. (8) [46], [47].

$$\begin{aligned} c_1^t = \mu * c_1^{t-1} * (1 - c_1^{t-1}) \\ c_2^t = \mu * c_2^{t-1} * (1 - c_2^{t-1}) \end{aligned} \tag{8}$$

where $\mu \in (0, 4]$ is the logistic function parameter, and the logistic chaotic function is in a completely chaotic state once the value of μ is taken as 4 [46], [47]. Secondly, the inertial weight ω is adaptively adjusted to balance the capabilities of exploitation and exploration, which can be calculated as Eq. (9) [30].

$$\omega^t = \omega_{max} - \frac{\omega_{max} - \omega_{min}}{t_{max}} \times (t - 1) \tag{9}$$

where t_{max} denotes the maximum number of iteration, ω_{max}, ω_{min} are the maximum and minimum inertia weights. Moreover, ω_{max}, ω_{min} are usually set to 0.9 and 0.4, respectively for the balance between global and local search [30], [48], [49]. Iteratedly, the positions and velocities of the particles are updated, and a fitness function is defined to calculate the values of the updated particles, obtaining the optimal $pbest_i$ and gbest in current rotation, which can be given in Eq. (10) [19], [30].

$$fitness = \frac{\sum_{i=1}^{k} r_{ui} \times \mu_i}{\sum_{i=1}^{k} \mu_i} \tag{10}$$

where r_{ui} is the output of rule i, k denotes the number of rules, and μ_i indicates the centroid of the fuzzy output membership function. Obviously, the global optimal solution can be reached when the number of iterations equals the preset t_{max} or the new updated gbest is not greater than the old one any more. Then, the optimal rules for fuzzy clustering can be obtained by decoding gbest, which are listed as follows:

Rule 1: if E_{res} = very less, V_{nd} = small, D_{toc} = very near, then chance = low

Rule 2: if E_{res} = less, V_{nd} = small, D_{toc} = very near, then chance = low

Rule 3: if E_{res} = medium, V_{nd} = small, D_{toc} = very near, then chance=medium

Rule 4: if E_{res} = much, V_{nd} = small, D_{toc} = very near, then chance=high

Rule 5: if E_{res} = very much, V_{nd} = small, D_{toc} = very near, then chance=very high

Rule 6: if E_{res} = very less, V_{nd} = medium, D_{toc} = very near, then chance=low

Rule 7: if E_{res} = less, V_{nd} = medium, D_{toc} = very near, then chance=low

Rule 75: if E_{res} = very much, V_{nd} = big, D_{toc} = very far, then chance=medium

For the same reason, the rules for fuzzy routing can also be obtained in the end. The processes for fuzzy rules optimization are described in Algorithm 3.

Similar to [19] and [30], the population size *Np=30*. Decoding the corresponding *gbest*, the output *chance/prob*

Algorithm 3 Pseudo Code for Rules Optimization

- **Input**
- Population size=Np
- **Output**
- Fuzzy rule set
- **Begin**

Initialization of population
Initialize particles $\mathbf{p_i} = \{\mathbf{x_{i1}}, \mathbf{x_{i2}}, \ldots, \mathbf{x_{in}}\}$ using Eq. (2), n =number of
nodes for clustering, or else =number of CHs for routing
for $i = 1$ to Np do
  Calculate Fitness($\mathbf{p_i}$) using Eq.(10)
  pbest_i =Fitness($\mathbf{p_i}$)
end for
gbest = max(pbest_i) $1 \le i \le$ Np
while (t<t_{max}) do
  for $i = 1$ to Np do
    update ω using Eq. (9)
    update $\mathbf{p_i}$ using Eqs. (6) and (7)
    if (Fitness($\mathbf{p_i}$) > pbest_i) then
      pbest_i =Fitness($\mathbf{p_i}$)
    end if
    if (pbest_i > gbest) then
      gbest = pbest_i
    end if
  end for
end while

- **End**

for fuzzy clustering and fuzzy routing can be obtained, respectively. The overall workflow chart of the proposed NFCRP protocol is depicted in Fig. 5.

In the figure, r denotes the current round, and r_{max} is the maximum round.

D. TIME COMPLEXITY ANALYSIS

In NFCRP, the overall time complexity includes the time complexity of fuzzy clustering, fuzzy routing and fuzzy rules optimization. For CHs selection in fuzzy clustering, a FIS is used to determine the CHs, whose time complexity is $O(n \times n_{rule})$, where n is the number of sensor nodes, and n_{rule} denotes the number of fuzzy rules which equals 75 in NFCRP. In addition, during the process of clusters formation, m messages are broadcast by CHs to announce their status, and (n-m) acknowledge message are sent to the CHs by CMs, where m is the number of CHs. Accordingly, the time complexity of fuzzy clustering is $O(n \times n_{rule} + n)$. For fuzzy routing, the other FIS is used to find the optimal relay CH for each CH, so the time complexity is also $O(m \times n_{rule})$, where m is the number of CHs. Besides, m messages are needed to broadcast the relay CH information. Hence, the time complexity of fuzzy routing is $O(m \times n_{rule} + m)$. Moreover, PSO is adopted to tune the fuzzy rules in NFCRP, and the PSO has time complexity $O(k \times n_p)$, where k is the number of sensor nodes or CHs, and n_p is the population size. Then, the time complexity of fuzzy rules optimization is $O((n + m) \times n_p)$. Therefore, the overall time complexity of NFCRP is $O((n + m + 2) \times n_{rule} + (n + m) \times n_p)$. Generally, m, n_p, and n_{rule} are much less than n, thus, the time complexity of NFCRP is $O(n^2)$.

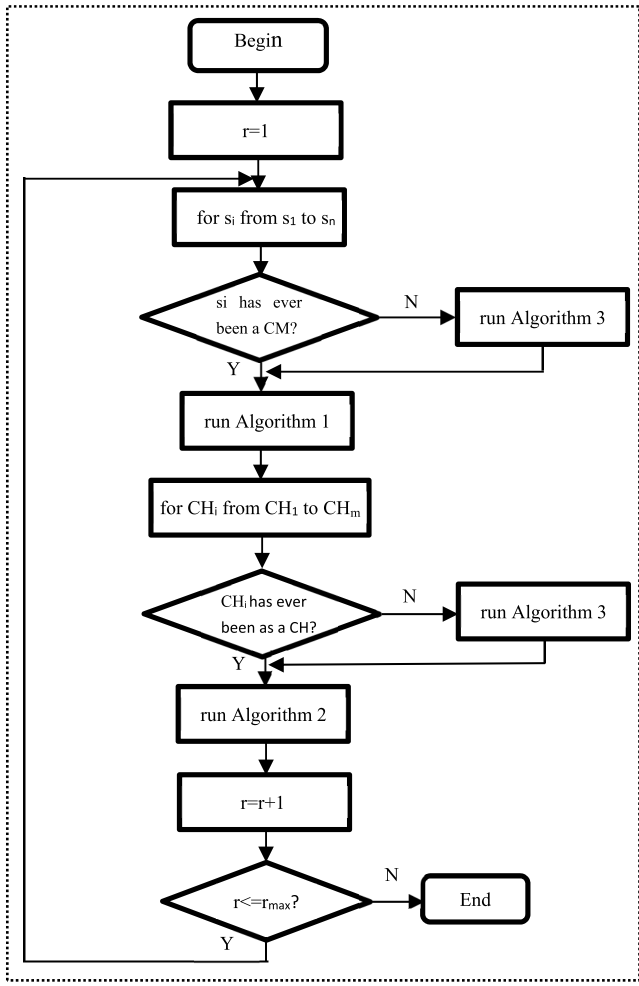


FIGURE 5. The overall workflow chart of the proposed NFCRP protocol.

V. SIMULATION RESULTS

In this section, the performance of the proposed NFCRP protocol is evaluated and compared with LEACH [3], EFUCA [14], EEFUC [16], FBCR [37] and FMSFLA [4]. The simulations are performed in MATLAB 2022a which is installed and operated over Windows 10 operating system with the hardware configuration of Intel Core, 4-GB RAM, i5-6300 CPU with a speed of 2.3 GHz. Fair comparisons are made by using the same parameters in all scenarios. Four simulation scenarios are considered and the parameters are given in the Table 3.

The results obtained in these four scenarios are averaged from 50 runs in each scenario. Four metrics network lifetime, standard deviation of CH's traffic load, network throughput and energy consumption are used for performance validation.

A. NETWORK LIFETIME

Firstly, network lifetime is tested in all scenarios. The network lifetime is directly related to the number of surviving nodes and its performance is usually measured in terms of

TABLE 3. Network parameters.

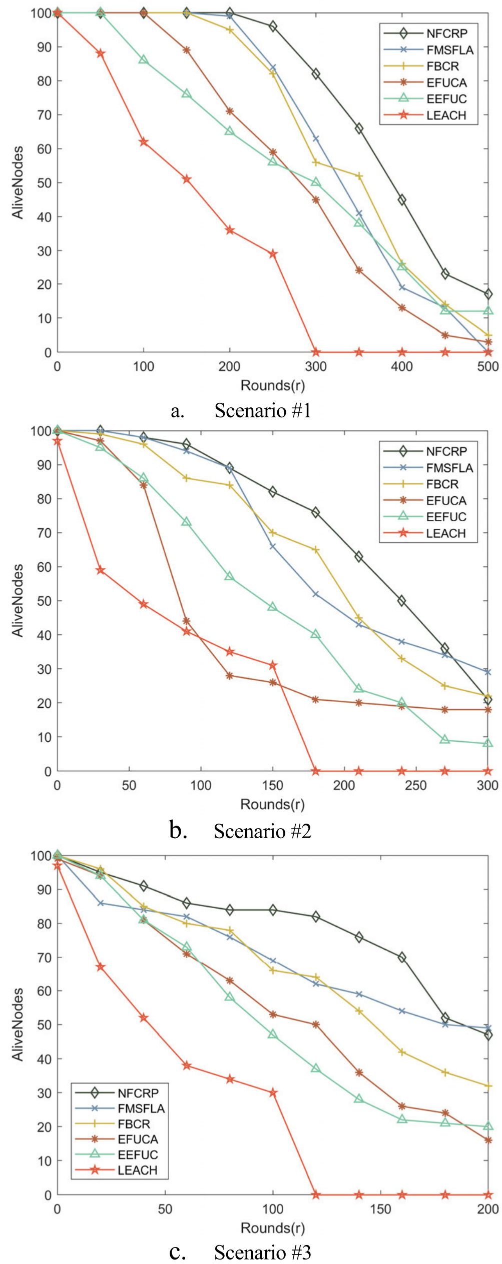
Parameters	Scenario #1	Scenario #2	Scenario #3	Scenario #4
Number of nodes	100	100	100	100
Initial energy	0.5J	0.5J	1J	1J
Eelec	50 (nJ/bit)	50 (nJ/bit)	50 (nJ/bit)	50 (nJ/bit)
E_0	5(nJ/bit)	5(nJ/bit)	5(nJ/bit)	5(nJ/bit)
ε_{fs}	10 (pJ/bit/m2)	10 (pJ/bit/m2)	10 (pJ/bit/m2)	10 (pJ/bit/m2)
ε_{mp}	0.0013 (pJ/bit/m4)	0.0013 (pJ/bit/m4)	0.0013 (pJ/bit/m4)	0.0013 (pJ/bit/m4)
d_0	87.7m	87.7m	87.7m	87.7m
Data packet size	4000bits	4000bits	4000bits	4000bits
Control packet size	200bits	200bits	200bits	200bits
Area	400m* 400m	400m* 400m	800m* 800m	800m* 800m
BS Location	x=200, y=200	x=200, y=400	x=400, y=400	x=400, y=800

FND, HND, and LND. The results of NFCRP compared to LEACH, EFUCA, EEFUC, FBCR and FMSFLA are shown in Table 4, Table 5 and Fig. 6.

It can be seen from Table 4 that NFCRP outperforms FMSFLA, EFUCA, EEFUC, FBCR and LEACH in terms of network lifetime. From scenario #1 to #4, the network lifetime of NFCRP is 14.67%, 17.83%, 8.67% and 11.00% higher than FMSFLA, 11.29%, 12.93%, 16.83% and 21.62% higher than FBCR, 33.89%, 61.53%, 55.10% and 50.89% higher than EFUCA, and 40.97%, 44.40%, 49.48% and 57.14% higher than EEFUC,71.66%, 79.37%, 78.06% and 89.28% higher than LEACH.

As can be seen from Table 5, in large-scale networks, the energy consumption of all the protocols increases because the inter-cluster distance also increases, which leads to a much shorter network lifetime. However, even in large-scale networks, the proposed NFCRP protocol in this paper slightly outperforms the FMSFLA, EFUCA, EEFUC, FBCR, and LEACH protocols. Therefor, NFCRP also has better scalability.

As can be seen from Fig.6, FBCR uses a single FIS for CHs selection using network partitioning and three indicators as fuzzy inputs, but using single hop for data transmission reduces network lifetime. Although both EFUCA and EEFUC use multiple FISs to select the best CHs and the optimal relay CHs based on different input parameters, however, nodes that are closer to the BS are prone to be elected as relays multiple times, resulting in premature death due to being overloaded. To overcome the problem of node's early death due to the low energy of the selected CHs, FMSFLA selects nodes with residual energy greater than the average energy of the network to participate in CH selection and



a. Scenario #1

b. Scenario #2

c. Scenario #3

FIGURE 6. Comparison of the number of alive nodes.

utilizes an optimization algorithm to adjust the fuzzy rules in CH selection. However, the traffic load on the parent nodes in FMSFLA is high, which affects the network lifetime. NFCRP overcomes the shortcomings of the other protocols and prolongs the network lifetime, so the quantity of the network survival nodes is generally more than that of other protocols.

B. STANDARD DEVIATION OF CH'S TRAFFIC LOAD

To verify the performance of load balancing, the standard deviation of CH's traffic load is tested due to the almost identical loads of the cluster members. Additionally, LEACH

TABLE 4. Comparison of FND, HND and LND.

		NFCRP	FMSFLA	FBCR	EFUCA	EEFUC	LEACH
Scenario #1	FND	206	182	176	105	54	17
	HND	387	324	350	287	296	151
	LND	637	491	586	512	543	298
Scenario #2	FND	49	51	36	23	10	1
	HND	237	184	213	87	149	58
	LND	459	423	413	381	365	176
Scenario #3	FND	8	2	1	1	8	1
	HND	188	177	162	117	91	42
	LND	537	502	463	402	434	139
Scenario #4	FND	3	2	1	1	2	1
	HND	108	98	86	54	46	11
	LND	306	260	226	161	187	84

TABLE 5. FND, HND and LND in the area of 1000m*1000m with different nodes.

		NFCRP	FMSFLA	FBCR	EFUCA	EEFUC	LEACH
600 Nodes	FND	37	21	24	16	10	1
	HND	171	156	161	146	142	17
	LND	481	372	413	347	363	217
1000 Nodes	FND	26	16	20	18	9	1
	HND	163	140	158	143	146	16
	LND	417	319	393	302	329	172

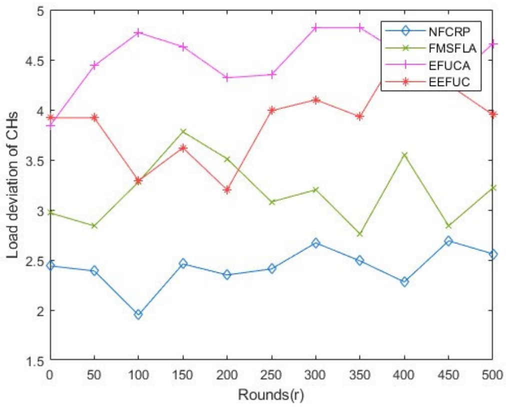
and FBCR transmit data from CHs to BS in single hop mode, so only the load deviation of the CHs of NFCRP, EFUCA, EEFUC, and FMSFLA are tested and the results are depicted in Fig.7.

As can be seen from Fig.7, in all scenarios, the standard deviation of CH's load for the NFCRP protocol is smaller than that of the FMSFLA, EFUCA and EEFUC protocols and remains stable. This is because NFCRP considers load balancing when finding routing paths and uses the number of times being selected as relay as a fuzzy input for relay node selection, thus mitigating the hot-spot problem. Due to its neglect of hot spot problem in EFUCA, it performs the worst in energy balance. Compared with EEFUC forming unequal clusters by adjusting the cluster's radius, FMSFLA considers the parameter the mean route load as a fuzzy input so as to balance the load of each route, resulting better energy balance performance. For NFCRP, data load deviation is considered as a fuzzy input to balance the inter-cluster energy consumption. Besides, the balanced intra-cluster energy consumption is also considered. Therefore, it performs best in energy balance. The experimental results show that the mean value of the load standard deviation of CHs for NFCRP is 36.7%, 30.57%, 34.93% and 14.96% lower than that of the EEFUC protocol, 45.32%, 30.94%, 36.17% and 13.27% lower than that of the EFUCA protocol, 21.57%, 16.13%, 44.09% and 19.34% lower than that of the FMSFLA protocol, respectively. Therefore, NFCRP is more effective in solving hot-spot problems.

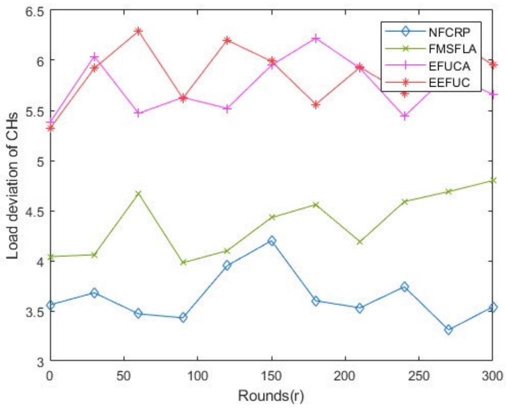
C. NETWORK THROUGHPUT

Next, the network throughput is tested in the four scenarios, which indicates the amount of data transmitted by nodes to the BS. A higher throughput indicates a higher utilization rate of energy in the network. The results are shown in Fig.8.

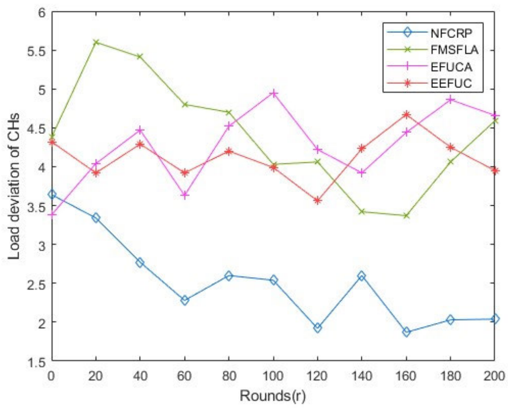
Observed from Fig.8, NFCRP has higher network throughput than LEACH, EFUCA, EEFUC, FBCR and FMSFLA in all scenarios. LEACH performs worst because of its single hop communication mode, resulting in premature death of the nodes far away from the BS. For FBCR, it performs much better than LEACH because it divides the network into four regions to avoid long distance transmission of the nodes, although it also adopts single hop communication mode. FMSFLA performs better than EEFUC and EFUCA because it can adjust its fuzzy rules. Moreover, NFCRP focuses on fuzzy rules themselves while FMSFLA aiming at the network lifetime for corresponding fitness function. However, the network lifetime is dynamic. Therefore, NFCRP forms appropriate CHs to forward more data to the BS than



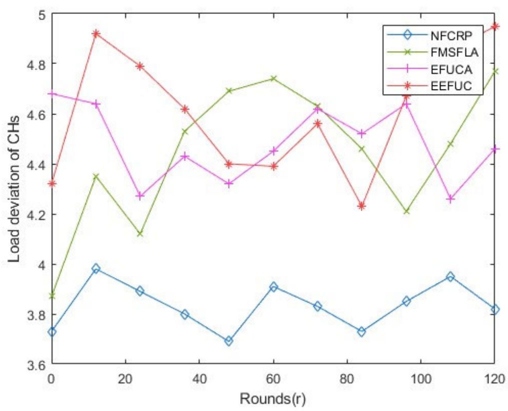
a. Scenario #1



b. Scenario #2

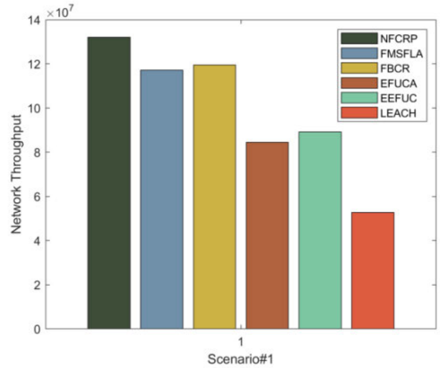


c. Scenario #3

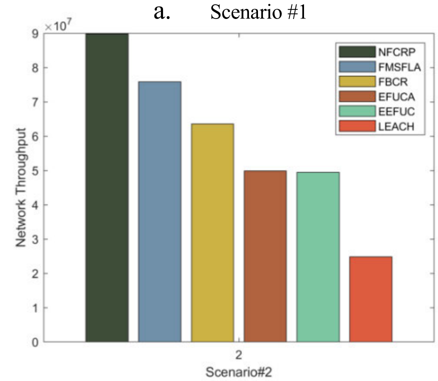


d. Scenario #4

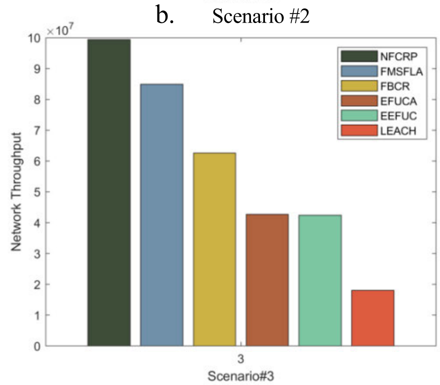
FIGURE 7. Comparison of standard deviation of CH's traffic load.



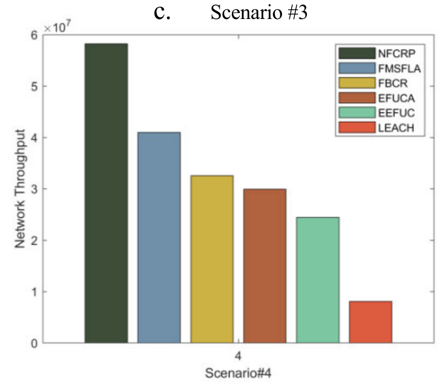
a. Scenario #1



b. Scenario #2



c. Scenario #3



d. Scenario #4

FIGURE 8. Comparison of network throughput.

FMSFLA. The network throughput of NFCRP is 60.13%, 70.55%, 79.45% and 86.06% higher than that of LEACH, 32.44%,45.07%, 57.26% and 57.96% higher than that of EEFUC, and 36.06%, 44.42%, 57.02% and 48.58% higher

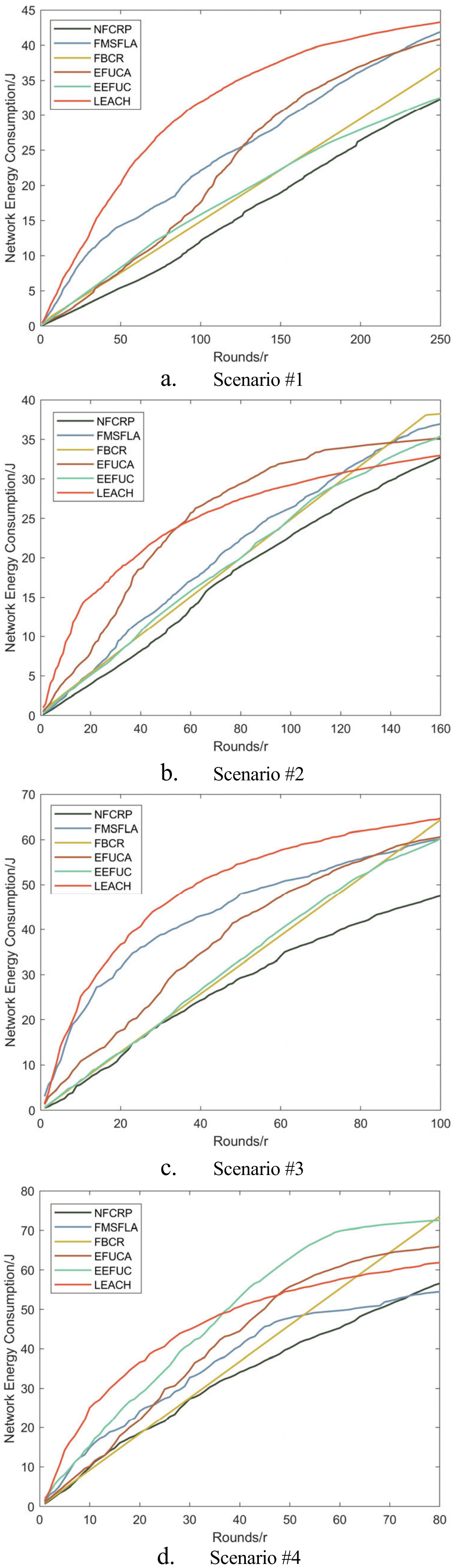


FIGURE 9. Comparison of the network energy consumption.

than that of EFUCA, 9.52%,29.27%, 37.01% and 44.08% higher than that of FBCR, and 11.35%, 17.20%, 9.29% and 29.65% higher than that of FMSFLA, respectively. Obviously, NFCRP not only prolongs the network lifetime, but also increases the amount of data transferred and further improves the network energy efficiency.

D. ENERGY CONSUMPTION

The total energy consumption in the network is tested to validate the overall performance of the network. The less the total energy consumption, the better the network performance. The results are displayed in Fig.9.

As can be seen from Fig. 9, the network energy consumption increases with the number of rounds, and the network energy consumption curve for NFCRP remains largely below the other protocols. The using of single hop mode for data transmission in FBCR affects the network lifetime to some extent, although it is better than LEACH with single hop mode as well. Both EFUCA and EEFUC select the next hop CH during data transmission based on the fuzzy logic output, which reduces network energy consumption to some extent. However, both the protocols make nodes close to the BS or with greater residual energy be selected as relays, which often results in closer, more energetic nodes being repeatedly selected as relays causing premature death. However, EEFUC can form unequal clusters by using a FIS to save intra-energy consumption. In addition, FMSFLA also reduces intra- and inter-cluster energy consumption by two FISs with adjustable rules. NFCRP aims to minimize the total energy consumption of the network by balancing energy consumption during the processes of clustering and routing. Hence, NFCRP runs 65.80%, 44.64%, 64.22% and 41.17% more rounds than LEACH, 10.88%, 10.71%, 29.35% and 44.11% more rounds than EEFUC, 35.23%, 47.32%, 38.53% and 33.82% more rounds than EFUCA, 11.91%, 9.82%, 28.44% and 12.69% more rounds than FBCR, 36.78%, 18.75%, 45.87% and 7.35% more than FMSFLA, respectively.

VI. CONCLUSION

A novel fuzzy clustering and routing protocol NFCRP is presented in this paper, which is not only energy-efficient, but also energy-balanced through managing the network energy dissipation effectively. In NFCRP, the fuzzy rules are optimized by improved particle swarm optimization algorithm for specific applications. The proposed protocol considers reasonable parameters including residual energy, node degree deviation, distance to centrality for clustering, and residual energy, distance to BS, data load deviation for routing with the purpose of achieving the global optimal solution. According to the simulation results, the NFCRP protocol outperforms LEACH, EEFUC, EFUCA, FBCR and FMSFLA in all scenarios with respect to network lifetime, standard deviation of CH's traffic load, network throughput and energy consumption. Specifically, the average network lifetime of NFCRP increased by 79.59%, 47.99%, 50.35%, 15.66%, 13.04%, compared to LEACH, EEFUC, EFUCA, FBCR and FMSFLA. For the average standard deviation of CH's traffic load, NFCRP decreased it by 29.29% over EEFUC, 31.42% over EAUCA, and 25.28% over FMSFLA. For network throughput, NFCRP outperformed LEACH, EEFUC, EFUCA, FBCR and FMSFLA by 16.87%, 46.52%, 48.18%, 29.97%, 71.79%. Finally, NFCRP also reduced energy consumption by 53.95%, 23.76%, 38.72%,15.71%, 27.18% as compared to LEACH, EEFUC, EFUCA, FBCR and FMSFLA, respectively. In a word, the proper formation of clusters and selection of next-hop CHs are simultaneously guaranteed in the proposed NFCRP protocol with the carefully considered parameters along with the optimized fuzzy rules.

Although NFCRP achieves good results in terms of network lifetime, throughput, energy consumption, it does not consider the security and robustness related factors in building reliable and secure clusters and routing paths. Moreover, the mobility of the nodes is not addressed, and the tests are performed only in the ideal network model. Therefore, for the future research, reliability and security among all the nodes in the network will be strived for. Furthermore, the mobile features of the BS will be taken into consideration as well. Finally, practical scenarios will be used to test the proposed protocol for applicability verification.

REFERENCES

[1] D. Kandris, C. Nakas, D. Vomvas, and G. Koulouras, "Applications of wireless sensor networks: An up-to-date survey," *Appl. Syst. Innov.*, vol. 3, no. 1, pp. 1–24, 2020.

[2] C. Del-Valle-Soto, A. Rodríguez, and C. R. Ascencio-Piña, "A survey of energy-efficient clustering routing protocols for wireless sensor networks based on Metaheuristic approaches," *Artif. Intell. Rev.*, vol. 56, no. 9, pp. 9699–9770, Sep. 2023.

[3] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Annu. Hawaii Int. Conf. Syst. Sci.*, Piscataway, NJ, USA, 2000, p. 10.

[4] F. Fanian and M. Kuchaki Rafsanjani, "A new fuzzy multi-hop clustering protocol with automatic rule tuning for wireless sensor networks," *Appl. Soft Comput.*, vol. 89, Apr. 2020, Art. no. 106115.

[5] J. Hou, J. Qiao, and X. Han, "Energy-saving clustering routing protocol for wireless sensor networks using fuzzy inference," *IEEE Sensors J.*, vol. 22, no. 3, pp. 2845–2857, Feb. 2022.

[6] Gunjan, A. K. Sharma, and K. Verma, "GA-UCR: Genetic algorithm based unequal clustering and routing protocol for wireless sensor networks," *Wireless Pers. Commun.*, vol. 128, no. 1, pp. 537–558, Jan. 2023.

[7] R. S. Y. Elhabyan and M. C. E. Yagoub, "Two-tier particle swarm optimization protocol for clustering and routing in wireless sensor network," *J. Netw. Comput. Appl.*, vol. 52, pp. 116–128, Jun. 2015.

[8] A. Giri, S. Dutta, and S. Neogy, "An optimized fuzzy clustering algorithm for wireless sensor networks," *Wireless Pers. Commun.*, vol. 126, no. 3, pp. 2731–2751, Oct. 2022.

[9] B. Bhushan and G. Sahoo, "FLEAC: Fuzzy logic-based energy adequate clustering protocol for wireless sensor networks using improved grasshopper optimization algorithm," *Wireless Pers. Commun.*, vol. 124, no. 1, pp. 573–606, May 2022.

[10] K. Arutchelvan, R. S. Priya, and C. Bhuvaneswari, "Honey badger algorithm based clustering with routing protocol for wireless sensor networks," *Intell. Autom. Soft Comput.*, vol. 35, no. 3, pp. 3199–3212, 2023.

[11] P. Joshi, S. Kumar, and A. S. Raghuvanshi, "A performance efficient joint clustering and routing approach for heterogeneous wireless sensor networks," *Expert Syst.*, vol. 40, no. 5, pp. 1–26, Jun. 2023.

[12] R. K. Yadav and R. P. Mahapatra, "Energy aware optimized clustering for hierarchical routing in wireless sensor network," *Comput. Sci. Rev.*, vol. 41, Aug. 2021, Art. no. 100417.

[13] M. Priti and G. Kapil, "HFLBSC: Heuristic and fuzzy based load balanced scalable clustering algorithm for wireless sensor network," *Wireless Pers. Commun.*, vol. 125, pp. 281–304, Jan. 2022.

[14] P. S. Mehra, "E-FUCA: Enhancement in fuzzy unequal clustering and routing for sustainable wireless sensor network," *Complex Intell. Syst.*, vol. 8, no. 1, pp. 393–412, Feb. 2022.

[15] C.-H. Wang, H.-S. Hu, Z.-G. Zhang, Y.-X. Guo, and J.-F. Zhang, "Distributed energy-efficient clustering routing protocol for wireless sensor networks using affinity propagation and fuzzy logic," *Soft Comput.*, vol. 26, no. 15, pp. 7143–7158, Aug. 2022.

[16] S. Phoemphon, C. So-In, P. Aimtongkham, and T. G. Nguyen, "An energy-efficient fuzzy-based scheme for unequal multihop clustering in wireless sensor networks," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 1, pp. 873–895, Jan. 2021.

[17] K. K. Le-Ngoc, Q. T. Tho, T. H. Bui, A. M. Rahmani, and M. Hosseinzadeh, "Optimized fuzzy clustering in wireless sensor networks using improved squirrel search algorithm," *Fuzzy Sets Syst.*, vol. 438, pp. 121–147, Jun. 2022.

[18] M. Shokouhifar and A. Jalali, "Optimized sugeno fuzzy clustering algorithm for wireless sensor networks," *Eng. Appl. Artif. Intell.*, vol. 60, pp. 16–25, Apr. 2017.

[19] A. Lipare, D. R. Edla, and R. Dharavath, "Fuzzy rule generation using modified PSO for clustering in wireless sensor networks," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 2, pp. 846–857, Jun. 2021.

[20] N. Ghosh, I. Banerjee, and R. S. Sherratt, "On-demand fuzzy clustering and ant-colony optimisation based mobile data collection in wireless sensor network," *Wireless Netw.*, vol. 25, no. 4, pp. 1829–1845, May 2019.

[21] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. Int. Conf. Neural Netw. (ICNN)*, vol. 4, 1995, pp. 1942–1948.

[22] V. Prakash and S. Pandey, "Metaheuristic algorithm for energy efficient clustering scheme in wireless sensor networks," *Microprocessors Microsyst.*, vol. 101, Sep. 2023, Art. no. 104898.

[23] M. Z. Ghawy, G. A. Amran, H. AlSalman, E. Ghaleb, J. Khan, A. A. Al-Bakhrani, A. M. Alziadi, A. Ali, and S. S. Ullah, "An effective wireless sensor network routing protocol based on particle swarm optimization algorithm," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–13, May 2022.

[24] Y. Xiuwu, Z. Zixiang, P. Wei, and L. Yong, "A novel multi-hop clustering routing algorithm based on particle swarm optimization for wireless sensors networks," *Wireless Pers. Commun.*, vol. 130, no. 2, pp. 935–956, May 2023.

[25] S.-H. Lee, C.-H. Cheng, C.-C. Lin, and Y.-F. Huang, "PSO-based target localization and tracking in wireless sensor networks," *Electronics*, vol. 12, no. 4, pp. 1–21, 2023.

[26] P. Rawat and S. Chauhan, "Particle swarm optimization based sleep scheduling and clustering protocol in wireless sensor network," *Peer-Peer Netw. Appl.*, vol. 15, no. 3, pp. 1417–1436, May 2022.

[27] M. Otair, O. T. Ibrahim, L. Abualigah, M. Altalhi, and P. Sumari, "An enhanced grey wolf optimizer based particle swarm optimizer for intrusion detection system in wireless sensor networks," *Wireless Netw.*, vol. 28, no. 2, pp. 721–744, Feb. 2022.

[28] X. Yang, X. Chen, R. Xia, and Z. Qian, "Wireless sensor network congestion control based on standard particle swarm optimization and single neuron PID," *Sensors*, vol. 18, no. 4, pp. 1–17, 2018.

[29] Y. X. Cao and J. B. Xu, "DV-hop-based localization algorithm using optimum anchor nodes subsets for wireless sensor networks," *Ad Hoc Netw.*, vol. 139, pp. 1–11, Feb. 2023.

[30] C. Wang, "A distributed particle-swarm-optimization-based fuzzy clustering protocol for wireless sensor networks," *Sensors*, vol. 23, no. 15, pp. 1–20, 2023.

[31] S. Choudhary, S. Sugumaran, A. Belazi, and A. A. A. El-Latif, "Linearly decreasing inertia weight PSO and improved weight factor-based clustering algorithm for wireless sensor networks," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 6, pp. 6661–6679, Jun. 2023.

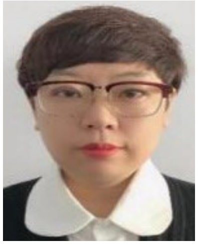
[32] I. Daanoune, B. Abdennaceur, and A. Ballouk, "A comprehensive survey on LEACH-based clustering routing protocols in wireless sensor networks," *Ad Hoc Netw.*, vol. 114, pp. 1–21, Apr. 2021.

[33] D. Kandris, E. A. Evangelakos, D. Rountos, G. Tselikis, and E. Anastasiadis, "LEACH-based hierarchical energy efficient routing in wireless sensor networks," *AEU-Int. J. Electron. Commun.*, vol. 169, Sep. 2023, Art. no. 154758.

[34] S. Z. Beevi and A. Alabdulatif, "Optimal routing protocol for wireless sensor network using genetic fuzzy logic system," *Comput., Mater. Continua*, vol. 70, no. 2, pp. 4107–4122, 2022.
[35] V. Sandeep, B. Sakshi, Z. Sherali, and S. Kaur, "Fuzzy-based techniques for clustering in wireless sensor networks (WSNs): Recent advances, challenges, and future directions," *Int. J. Commun. Syst.*, vol. 8, pp. 1–25, Nov. 2023.
[36] J.-S. Lee and W.-L. Cheng, "Fuzzy-logic-based clustering approach for wireless sensor networks using energy predication," *IEEE Sensors J.*, vol. 12, no. 9, pp. 2891–2897, Sep. 2012.
[37] A. Naderloo, S. A. F. Aghda, and M. Mirfakhraei, "Fuzzy-based cluster routing in wireless sensor network," *Soft Comput.*, vol. 27, no. 10, pp. 6151–6158, May 2023.
[38] M. Gamal, N. E. Mekky, H. H. Soliman, and N. A. Hikal, "Enhancing the lifetime of wireless sensor networks using fuzzy logic LEACH technique-based particle swarm optimization," *IEEE Access*, vol. 10, pp. 36935–36948, 2022.
[39] S. A. Sert, H. Bagci, and A. Yazici, "MOFCA: Multi-objective fuzzy clustering algorithm for wireless sensor networks," *Appl. Soft Comput.*, vol. 30, pp. 151–165, May 2015.
[40] Y. H. Robinson, E. G. Julie, R. Kumar, and L. H. Son, "Probability-based cluster head selection and fuzzy multipath routing for prolonging lifetime of wireless sensor networks," *Peer-Peer Netw. Appl.*, vol. 12, no. 5, pp. 1061–1075, Sep. 2019.
[41] V. Narayan, A. K. Daniel, and P. Chaturvedi, "E-FEERP: Enhanced fuzzy based energy efficient routing protocol for wireless sensor network," *Wireless Pers. Commun.*, vol. 131, no. 1, pp. 371–398, Jul. 2023.
[42] Z. M. Zahedi, R. Akbari, M. Shokouhifar, F. Safaei, and A. Jalali, "Swarm intelligence based fuzzy routing protocol for clustered wireless sensor networks," *Expert Syst. Appl.*, vol. 55, pp. 313–328, Aug. 2016.
[43] S. Janakiraman, "Improved bat optimization algorithm and enhanced artificial bee colony-based cluster routing scheme for extending network lifetime in wireless sensor networks," *Int. J. Commun. Syst.*, vol. 36, no. 5, pp. 1–20, Mar. 2023.
[44] P. Rawat, P. Kumar, and S. Chauhan, "Fuzzy logic and particle swarm optimization-based clustering protocol in wireless sensor network," *Soft Comput.*, vol. 27, no. 9, pp. 5177–5193, May 2023.
[45] Y. Zhang, J. Wang, D. Han, H. Wu, and R. Zhou, "Fuzzy-logic based distributed energy-efficient clustering algorithm for wireless sensor networks," *Sensors*, vol. 17, no. 7, pp. 1–21, Jul. 2017.
[46] Y. Peng, K. Lei, X. Yang, and J. Peng, "Improved chaotic quantum-behaved particle swarm optimization algorithm for fuzzy neural network and its application," *Math. Problems Eng.*, vol. 2020, pp. 1–11, Mar. 2020.
[47] L.-H. Ye, S.-J. Chen, Y.-F. Shi, D.-H. Peng, and A.-P. Shi, "Remaining useful life prediction of lithium-ion battery based on chaotic particle swarm optimization and particle filter," *Int. J. Electrochem. Sci.*, vol. 18, no. 5, May 2023, Art. no. 100122.
[48] R. Karimi, V. Shokri, M. Khishe, and M. K. Jameie, "Marine propeller design using evolving chaotic autonomous particle swarm optimization," *Wireless Pers. Commun.*, vol. 125, no. 2, pp. 1653–1675, Jul. 2022.
[49] Y. Duan, N. Chen, L. Chang, Y. Ni, S. V. N. S. Kumar, and P. Zhang, "CAPSO: Chaos adaptive particle swarm optimization algorithm," *IEEE Access*, vol. 10, pp. 29393–29405, 2022.

LIU YUEBO received the B.Sc. and M.S. degrees from the College of Computer Science and Engineering, Changchun University of Technology, in 2003 and 2006, respectively. Since 2006, she has been the Vice Professor with the College of Computer Engineering and Artificial Intelligence, Jilin University of Architecture and Technology. She has more than 50 research publications and more than 20 inventions. Her current research interests include wireless sensor networks, big data, data mining, and image processing.

YU HAITAO received the B.Sc. and M.S. degrees from the College of Computer Science and Technology, Northeast Normal University, in 2010 and 2013, respectively. Since 2013, he has been an Engineer with the Department of Information Center, Jilin Communications Polytechnic. He has more than 30 research publications. His current research interests include wireless sensor networks, big data, mobile sensing security, and image processing.



LI HONGYAN received the B.Sc. and M.S. degrees from the College of Business Administration, Northeast Normal University, in 2011 and 2014, respectively. Since 2014, she has been the Vice Professor with the School of Management, Changchun Information Technology College. She has more than 30 research publications and more than ten inventions. Her current research interests include wireless sensor networks, big data, data mining, and image processing.

LIU QINGXUE received the B.Sc. and M.S. degrees from the College of Computer Science and Technology, Northeast Normal University, in 2004 and 2007, respectively. Since 2007, she has been a Professor with the College of Computer Engineering and Artificial Intelligence, Jilin University of Architecture and Technology. She has more than 50 research publications and more than 20 inventions. Her current research interests include wireless sensor networks, signal and information processing, big data, and image processing.