Received November 9, 2019, accepted November 20, 2019, date of publication November 26, 2019,
date of current version December 9, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2955993

Energy-Efficient Routing in WSN: A Centralized Cluster-Based Approach via Grey Wolf Optimizer

S. M. MAHDI H. DANESHVAR, PARDIS ALIKHAH AHARI MOHAJER, AND SAYYED MAJID MAZINANI
Department of Computer and Electrical Engineering, Imam Reza International University, Mashhad 553-91735, Iran
Corresponding author: Sayyed Majid Mazinani (smajidmazinani@imamreza.ac.ir)

ABSTRACT Energy efficiency is one of the main challenges in developing Wireless Sensor Networks (WSNs). Since communication has the largest share in energy consumption, efficient routing is an effective solution to this problem. Hierarchical clustering algorithms are a common approach to routing. This technique splits nodes into groups in order to avoid long-range communication which is delegated to the cluster head (CH). In this paper, we present a new clustering algorithm that selects CHs using the grey wolf optimizer (GWO). GWO is a recent swarm intelligence algorithm based on the behavior of grey wolves that shows impressive characteristics and competitive results. To select CHs, the solutions are rated based on the predicted energy consumption and current residual energy of each node. In order to improve energy efficiency, the proposed protocol uses the same clustering in multiple consecutive rounds. This allows the protocol to save the energy that would be required to reform the clustering. We also present a new dual-hop routing algorithm for CHs that are far from the base station and prove that the presented method ensures minimum and most balanced energy consumption while remaining nodes use single-hop communication. The performance of the protocol is evaluated in several different scenarios and it is shown that the proposed protocol improves network lifetime in comparison to a number of recent similar protocols.

INDEX TERMS Clustering, grey wolf optimizer, routing, WSN.

I. INTRODUCTION

Wireless sensor networks (WSNs) are emerging low-cost and versatile solutions that enable controlled monitoring of the environment. They generally consist of a large number of small sensing devices that are capable of data processing and wireless communication. These sensor nodes can be deployed in various environments to implement applications such as habitat monitoring, military surveillance, home and industrial automation, and smart grids [1], [2]. Recent advances in electronic circuit design have made it possible to build lighter, cheaper and more energy efficient sensors. However many research areas including energy efficiency need to be further studied [3]. In many applications, sensor nodes are equipped with a non-rechargeable battery that restricts network lifetime [4]. There are several definitions for lifetime, such as the time until the first node dies or the time that the last node dies or the time until a specific fraction of nodes die [5]. After the death of the first node, the performance of the network will degrade sharply [6]. In [7] and [8], network lifetime is defined in terms of node lifetime, coverage, and connectivity. Although the use of renewable energy sources for sensor nodes are investigated in Energy Harvesting Wireless Sensors Networks (EHWSN) [9], wise use of the available energy is still required for long running WSNs. Most WSNs measure physical parameters such as temperature, humidity or location of objects. Samples of these parameters are locally correlated, therefore can be aggregated for neighbor sensors. The energy required for data transmission is several hundred times greater than the energy required for data processing [10]. Therefore it's wise to compress data before transmission. This data compression via signal aggregation leads to a significant reduction in energy spent on communication and prolongs the network lifetime [11]. Hierarchical clustering protocols use this fact to extend network lifetime by splitting nodes into several spatial clusters in which, only one of the sensors (CH)

The associate editor coordinating the review of this manuscript and approving it for publication was Peng-Yong Kong.

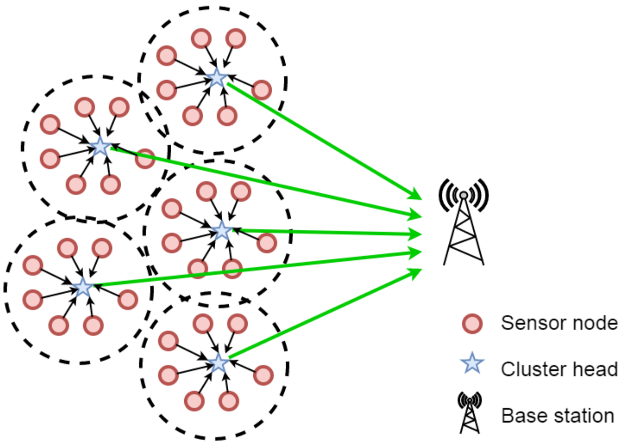

FIGURE 1. **Structure of a typical WSN.**

is responsible for aggregating the signals and sending data to the base station (BS). All the other nodes engage only in intra-cluster short-range communication which consumes much less energy. A CH can either send the data directly to the BS (single-hop) or use another node as a relay to deliver the data to the BS (multi-hop) [12]. The clustering protocols also divide lifetime of the network into several rounds and try to rotate the role of CH among all nodes in different rounds to achieve balanced energy consumption. Fig. 1 shows a typical structure of a hierarchical clustering over a WSN.

Clustering has several advantages: (1) It reduces long range communications and overall energy consumption, (2) It reduces channel contention and packet collision and (3) It results in better throughput under high load.

The choice of CHs has a significant effect on the performance of a protocol. CH election is an NP-hard problem because the selection of m optimal CHs out of n nodes gives $\binom{n}{m}$ possibilities [13], Thus classical optimization algorithms are not feasible [14]. Because of this, many clustering algorithms rely on heuristic and metaheuristic approaches to select relatively optimal CHs to different degrees of success. grey wolf optimizer (GWO) [15] is one of the recent metaheuristic swarm intelligence methods. It's inspired by the behavior of grey wolves when hunting prey. Compared to other swam intelligence methods, it has fewer parameters. It also balances its search such that both exploration and exploitation are achieved. It gives the algorithm the ability to better escape local-optimum which leads to improved convergence [16].

In this paper, we propose a new clustering and routing protocol which uses GWO to select optimal CHs. Clustering is done at the BS, therefore the presented protocol is centralized. It chooses a set of CHs that split the network into several clusters. To avoid fast energy depletion of CHs that are far from BS, a relay node is selected for each of them. Many clustering protocols divide each round into two phases: (a) cluster setup phase and (b) steady state phase. In the setup phase the clusters are formed and in the steady state phase nodes send their data to the BS via CHs. Forming clusters poses an energy overhead over the network as it leads to consumption of a certain amount of energy due to exchange of necessary control packets [17]. The proposed protocol minimizes this overhead by eliminating the setup phase in some rounds where the CHs from previous rounds are still good enough. In other words, BS trusts the CHs of the previous round and skips the setup phase in some of the rounds. This continues until at least one of CHs has consumed 50% of its energy since the last setup phase in which case the setup phase is executed again in the beginning of the next round to select new CHs. Otherwise, the network continues to operate without executing the setup phase.

The rest of this paper is structured as follows. We review related research in section I-A, and present a summary of our contributions in section I-B. The system model is presented in section II. We discuss the proposed protocol in section III and how to apply gray wolf optimizer to the problem in section IV. Section V deals with protocol analysis and experimental results are presented in section VI.

A. RELATED WORK

Designing energy efficient clustering algorithms has been subject of intensive research recently. One of the most well-known clustering protocols is the LEACH [18] which selects CHs based on a predetermined probability in order to rotate CH role among sensors by using a random variable. Although LEACH balances energy consumption in the network, it has several shortcomings: i) it does not take energy of the nodes into account. Therefore it is possible that nodes with lowest remaining energy become CHs, ii) it does not take location of CHs into account and iii) it uses single-hop communication which results in fast energy depletion of CHs that are far from the BS. LEACH-C [19], is a centralized version of LEACH. It uses simulated annealing to minimize inter-cluster distance. A probability proportional to energy of the node is used to select CHs. LEACH-C manages to further improve the lifetime of the network but is unable to form the best possible clusters due to its random nature. In addition to this, using a single hop routing method causes more distant nodes from the BS to die much faster. This results in drastic loss of coverage and shortening of network lifetime. PSO-ECHS [14] is a clustering protocol that aims to prolong network lifetime by using a centralized clustering algorithm based on PSO [20]. The PSO based algorithm finds a suitable subset of nodes for CHs along with their respected cluster members while trying to: i) minimize the average distance between cluster members and CHs, ii) minimize the average distance between BS and CHs, and iii) maximize sum of energy of CHs. To select a CH for each member, parameters such as energy, distance to the BS and node degree of the CH are taken into account. PSO-ECHS manages to improve network lifetime and total energy consumption of the network; however, it does not consider any routing algorithm and relies on single-hop communication which leads to early death of distant nodes to the BS. The protocol proposed in [11] uses an improved

PSO algorithm to select CHs while considering two metrics: i) ratio of total energy of CHs to total energy of members, ii) ratio of the average distance between non-CHs to the BS and the average distance between clusters heads to the BS. In addition, to prevent fast energy depletion of distant nodes which previously mentioned protocols suffered from, this protocol chooses relay nodes to offload energy consumption of CHs. A separate PSO-based algorithm is proposed to select suitable relay nodes for each of the CHs. The algorithm tries to maximize ratio of average energy of the relay nodes to the average energy of other nodes. Therefore more powerful nodes are selected as relay nodes. The other factor that is taken into account is the location of the relay node. The relay node which minimizes the transmission cost between CH and the BS is selected. This protocol is successful in extending network lifetime and solves many shortcomings of the previous protocols; however, there is no guarantee that the selected clustering is the best possible solution due to the nature of PSO. PSO-UCF [4] is another PSO-based protocol to prolong network lifetime. It splits the network to unequal clusters and uses multi-hop communication between CHs and the BS. To add fault tolerance and aid energy efficiency, the algorithm also selects a surrogate CH for each cluster. To choose CHs, the PSO algorithm is used in order to: i) minimize average intra-cluster communication distance, ii) minimize average inter-cluster communication distance and iii) maximize sum of energy of CHs. The closest cluster member to CH with enough energy to last for a round is selected as surrogate CH. After that, a multi-hop routing tree is formed by selecting next hop for each CH among other CHs. This selection is done by considering energy and distance. Each node then joins a CH with the maximum weight which takes distance to the sensor node, distance to the BS, node degree and residual energy into account. This results in formation of unequal clusters where clusters near the BS have more members than distant clusters. To further improve energy efficiency, this protocol avoids new cluster formation in beginning of each round; instead, new clusters are formed only when energy of some CHs falls below a calculated threshold which is an approximation of amount of energy needed for each node to operate for one round under current clustering. PSO-UFC manages to show very competitive results while addressing problems such as hot spot problem and fault tolerance however does not guarantee that best solution is found. SMOTECP [21] is a similar clustering protocol that uses Spider Monkey Optimizer [22] to form clusters by minimizing energy consumption and cluster cohesion while maximizing cluster separation and energy of CHs. CHs that are closer to the BS send their data directly to the BS but distant CHs choose a closer CH to the BS as their relay node. While SMOTECP extends network lifetime to some degree, it wastes energy on cluster formation as it is done in every round. The protocol presented in [23] assumes a mobile sink and divides network into two stages. In the first stage, a geometric method is used to select CHs. After a number of rounds, the energy of nodes becomes unequal and network goes to second stage. In the second stage, it uses an improved PSO algorithm to maximize energy of nodes that are close to CHs in order to avoid hot spot problem. Therefore it will choose nodes closer to energy centers as CHs. Every CH uses a greedy method to determine its next hop. This protocol executes CH selection every 20 rounds to minimize the energy overhead of exchanging control packets for CH selection. UCRA-GSO [24] uses glowworm swarm optimization to produce uneven clusters by taking cluster compactness and degree, energy, and proximity of CHs into account. It also forms a routing tree by selecting a best next hop for every CH, this selection is based on the communication distance and residual energy of the next hop. UCRA-GSO executes setup phase in every round which leads to an avoidable and unnecessary energy waste. MLHP [25] is a hybrid protocol that splits network into three spatial levels based on the distance form BS. It chooses CHs based on distance form BS and residual energy in the closest level to BS. In the second level a GWO-based algorithm is used to choose suitable CHs by presenting a probabilistic fitness function which takes residual energy and number of neighbors into account. The CH selection in the third level is distributed and a routing tree is constructed to deliver data to the BS.

B. CONTRIBUTIONS

The protocols reviewed in the previous section prolong the network lifetime to some extent. But there is still room for improvement. Energy consumption is directly related to distance between sender and receiver in transmissions. Therefore, all the aforementioned protocols aim to minimize energy consumption by minimizing distance of transmissions. In other words, using distance is an approximation to energy that will be required for communication [19]. However, given the fact that these protocols are centralized and the BS has unlimited energy and high computational power, it's possible to accurately compute the energy that the network would consume during the next round. Therefore instead of relying on the inter-cluster and intra-cluster distances to approximate energy consumption and assign fitness value to the solution, we introduce a single term fitness function that is directly related to the energy consumed by the network. This has two advantages: i) assigning fitness values to solutions more accurately, and ii) eliminating the need to weight several terms in the fitness function which leads to fewer parameters for the protocol. In addition, the proposed protocol enables the BS to approximate how many rounds each node can operate using the current clustering and use this information to skip cluster setup phase in some of the rounds. This leads to a significant reduction in energy consumption due to skipping the exchange of control packets. We also propose a novel solution to this optimization problem using the grey wolf optimizer and a suitable way to present the clustering problem as an optimization problem. The proposed protocol uses dual hop routing to avoid fast depletion of energy for CHs that are far from the BS. We prove that there exists a specific point between CH and BS which is the perfect point for a relay to both minimize and balance energy consumption of relay and CH. Using this definition we propose a method to select relay nodes. This ensures energy consumption is balanced for both far and close CHs to the BS while minimum energy is consumed. Our contribution can be summarized to:

1) Solving the WSN energy efficient clustering problem using the Gray Wolf Optimizer by introducing a fitness function based on the predicted energy consumption.
2) Proposing a method to skip setup phase in some rounds in order to eliminate corresponding energy overhead while keeping the network connected and efficient.
3) Proposing a dual hop routing method that ensures both minimum and most balanced energy consumption for any CH and its relay node.

II. SYSTEM MODEL

The proposed network model consists of a BS and N sensors that are deployed randomly with uniform distribution in the network area. In order to ensure network connectivity, large number of nodes are deployed in the area which prevents the network from dividing into several isolated areas [26]. The following assumptions are made about the network:

1) All sensor nodes and the BS remain stationary after deployment.
2) Sensor nodes are capable of reporting their location.
3) All nodes have equal initial energy.
4) The BS has high computational power and is not energy limited.
5) Sensor nodes measure the environment at a fixed rate and send data periodically to their CH.
6) The nodes are able to adjust their transmission power according to the distance to the target.
7) BS is within the communication range of all nodes.
8) The network is in a favorable environment, radio channel is symmetric with no collision during transmission [22].

The energy model used in this paper is the same as the one presented in [11]. In this model, either free space (d^2 power loss) or multipath fading (d^4 power loss) channel model is employed based on the distance between the sender and the receiver. If the distance is less than the threshold d_0, then the free space model is used, otherwise, the multipath fading model will be used. The energy consumed to transmit an l-bit message over the distance d is given by (1).

$$E_{TX}(l,d) = \begin{cases} l \times E_{elec} + l \times E_{fs} \times d^2, & d \le d_0 \\ l \times E_{elec} + l \times E_{mp} \times d^4, & d > d_0 \end{cases} \tag{1}$$

where E_{elec} is the amount of energy dissipated by the electronic circuit per bit. E_{fs} and E_{mp} depend on the transmitter amplifier model and d_0 is the distance threshold which is given by (2).

$$d_0 = \sqrt{\frac{E_{fs}}{E_{mp}}} \tag{2}$$

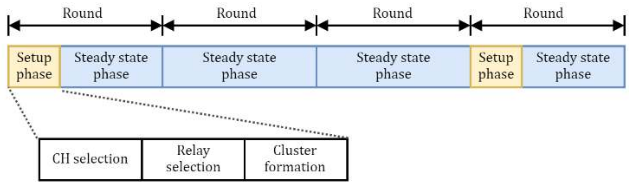

FIGURE 2. Overall operation of the protocol.

In addition, energy consumed to receive an l-bit message by the radio is given by (3).

$$E_{RX}(l) = l \times E_{elec} \tag{3}$$

III. PROPOSED PROTOCOL

In the proposed protocol, the network lifetime is divided into several rounds. The operation of protocol is in two phases: setup phase and steady state phase. In the setup phase, the BS gathers information including energy and location from nodes, BS then chooses CHs by using the proposed GWO based algorithm as well as a dual-hop route between CHs and BS. In the steady state phase, nodes send collected data to CHs. CHs aggregate and send data to the BS either directly or via another CH. To improve energy efficiency, this protocol executes setup phase only if current CHs are dying soon. This saves the energy that would be required to exchange control packets necessary to form clusters. Fig.2 shows the overall operation of the protocol.

A. CH SELECTION

Here, we represent the WSN clustering as an optimization problem. To this end, we introduce time of death T_D of a node as the number of rounds that the node can operate given the current clustering. This can also be formulated as (4).

$$T_D(i) = \frac{E_r(i)}{E_c(i)} \tag{4}$$

where $E_r(i)$ is the residual energy of the node and $E_c(i)$ is the amount of energy that the node consumes in a single round according to its role in the network. Nodes have different T_D values in different solutions and the best solution is the one that maximizes average T_D over all nodes. This can be formulated as an optimization problem expressed by (5).

$$maximize f = avg(\mathrm{T}_D) = \frac{1}{|nodes|} \sum_{i \in nodes} T_D(i) \tag{5}$$

This is subject to the following constraint:

$$E_r\left(CH_j\right) > \frac{1}{|nodes|} \sum_{i \in nodes} E_r(i), \quad 0 < j \le m$$

This states that the remaining energy of every CH should be greater than the average energy of all nodes. The reason to apply this constraint is to avoid solutions in which low-energy nodes are CH.

Time of death of each node can be calculated based on its role in the network. Here we did not take energy consumed by sensing and processing into account because they are insignificant compared to energy spent on communication [27]. The energy that a member node consumes in a single round is given by (6). The energy corresponding to exchange of control packets is not taken into account, because we are calculating how many rounds a node can operate under current clustering, we are implicitly assuming that no setup phase is executed therefore no control packet is exchanged.

$$E_c^{member}(i) = E_{TX}(L, dis(i, CH_i)) \tag{6}$$

Equation (6) corresponds to sending a message of length L to the CH. Note that L is a constant. More specifically, E_{TX} is the transmission energy given by (1). L is the length of the message in bits, $dis(i,j)$ is the distance between node i and node j and CH_i is CH of node i. The energy that a CH consumes in a single round is given by (7).

$$\begin{aligned} E_c^{ch}(j) = E_{RX}\left(L \times CM_j\right) + E_{DA} \times L \times (CM_j + 1) \\ + E_{TX}(L, dis(j, next(j))) \end{aligned} \tag{7}$$

where E_{RX} is reception energy given by (3), E_{DA} is data aggregation energy per bit; CM_j is the number of member nodes belonging to CH_j and *next* is the next hop which is either another CH or the BS. This corresponds to receiving and aggregating data from a number of members and sending the aggregated data to the next hop. Additionally, some CHs act as relay for one other CH, the energy consumed to relay data is given by (8). This corresponds to receiving an L bit message and sending it to the BS.

$$E_c^{relay}(r) = E_{RX}(L) + E_{TX}(L, dis(r, BS)) \tag{8}$$

The proposed protocol uses grey wolf optimizer to find an optimal solution for this problem. This process is described in section IV.

B. RELAY NODE SELECTION

Besides the selection of CHs, in order to prevent fast energy depletion of CHs which are far from BS, the protocol chooses a relay node for each of them in such a way that each relay is only used by one CH. If a relay is used by multiple CHs then the relay needs to assign as many timeslots to handle the communications which decreases the throughput of the network. Therefore in our protocol, a relay node is used by exactly one CH. As a result, some CH will not have a relay and send data directly to the BS. The algorithm to select relay nodes starts with the farthest CH from the BS and assigns a suitable relay node to it, and then it proceeds to run the same procedure for other CHs, starting with farther ones. To select suitable relay nodes, two goals are considered: i) minimizing total energy consumption and ii) balancing energy consumption between CHs and their relays. Fig.3 shows a CH, the BS and a hypothetical relay. Distance from CH to BS divided by r_0 is equal to the distance between CH to relay.

We will show in section V that if we use dual hop routing and one-to-one relation between CHs to relays, a specific fixed value for r_0 can be found which both minimizes and balances the energy consumption at the same time. In other

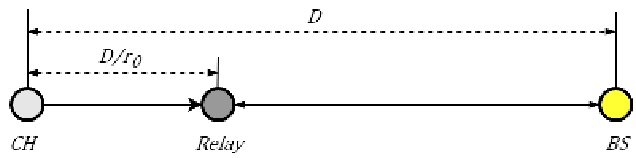

FIGURE 3. BS and a CH with location of a hypothetical relay node.

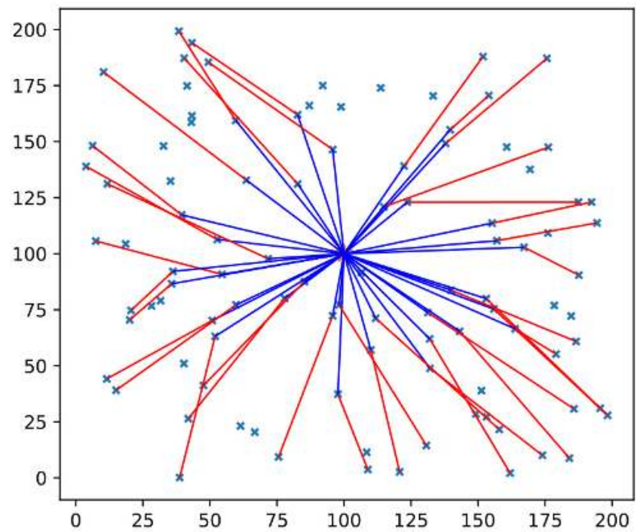

FIGURE 4. Result of relay selection. Red lines are from CH to relays and blue lines are from relays to the BS.

words, there is a point on the line segment between the CH and the BS that is the perfect point for the relay. This ensures minimum energy is consumed to deliver the packets to the BS and that CH and its relay consumed an equal amount of energy. We calculated $r_0 = 1.8169$. There will be a thorough explanation of how this value is calculated in section V.

To select a relay for each CH, the perfect point for the relay node between CH and BS is calculated based on the fixed distance ratio $r_0 = 1.8169$. Then the closest CH to that point, which is not already a relay, is selected. If no relay is closer than a threshold T_r to the perfect point, the CH will have no relay and sends directly to the BS.

This procedure has several advantages:

1) It avoids the hot spot problem because it ensures that a relay consumes an equal amount of energy to the CH it serves and that it selects different relays in different rounds.
2) It minimizes the total energy consumed to deliver data packets to BS
3) It tries to assign relays to as many CHs as possible.

An example of running this algorithm for relay selection is shown in Fig. 4. In this network, there are 100 CHs and BS is located at the center. Note that some CHs do not have a relay, which are depicted as a cross with no line.

C. CLUSTER FORMATION

At the beginning of the setup phase, each node sends a Node-MSG to the BS which contains the remaining energy

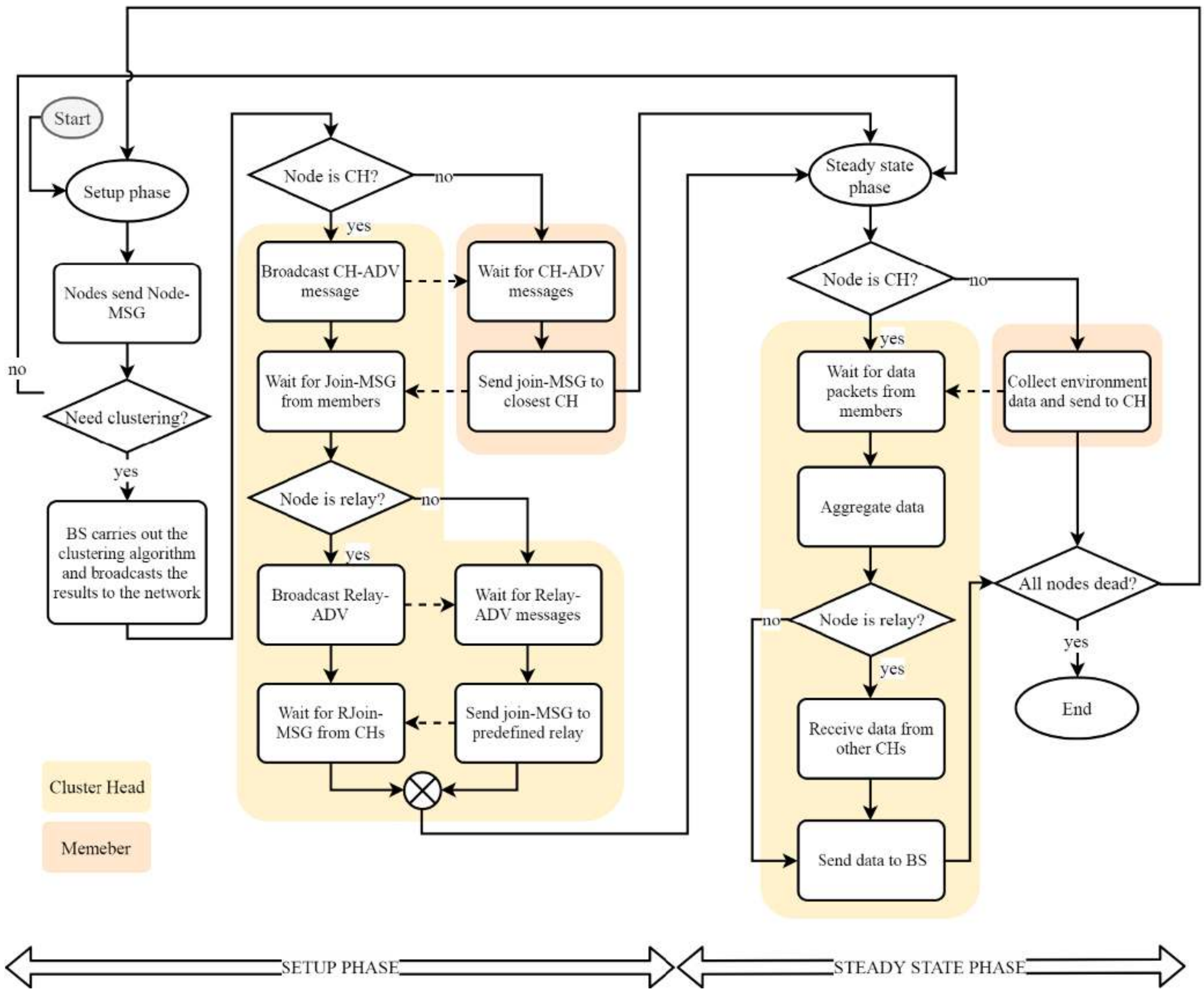

FIGURE 5. Flowchart of the protocol operation.

and the location of the node. These parameters are essential for BS to carry out the clustering algorithm. The BS then uses the proposed GWO algorithm to select a set of CHs that maximizes the fitness function given by (5). BS then broadcasts a message containing the ID of the CHs and relays to the network. After each CH is aware that it has been selected as CH, it broadcasts a CH-ADV message to introduce itself to the network. Member nodes choose the CH which requires minimum transmission energy according to the power of the received CH-ADV message and send a Join-MSG to the selected CH. Relay nodes then broadcast a Relay-ADV message. At this point, because CHs know the ID of their relay node they wait to receive Relay-ADV from their relay then send RJoin-MSG to the predefined relay node. After this procedure is complete, every node is aware of its role and the network can proceed to the steady state phase. The setup phase is not executed in the consecutive rounds until energy of one node falls below half of its energy since the last setup phase. The entire operation of the protocol is shown in Fig.5.

D. STEADY-STATE PHASE

The steady-state phase contains two stages:

1) Intra-cluster communication: Each node wakes up in its allocated time-slot and sends collected data directly to the corresponding CH.

2) Inter-cluster Communication: after all of the member nodes sent their data to the CHs, each CH aggregates the received cluster data then forwards the aggregated data to BS or its relay node. Each relay node forwards data of a single CH without aggregation. Therefore, relay nodes send exactly two data packets to the BS in each round.

IV. CLUSTERING BASED ON GREY WOLF OPTIMZER

In recent years many optimization algorithms are employed in clustering of WSNs. The swarm based optimization

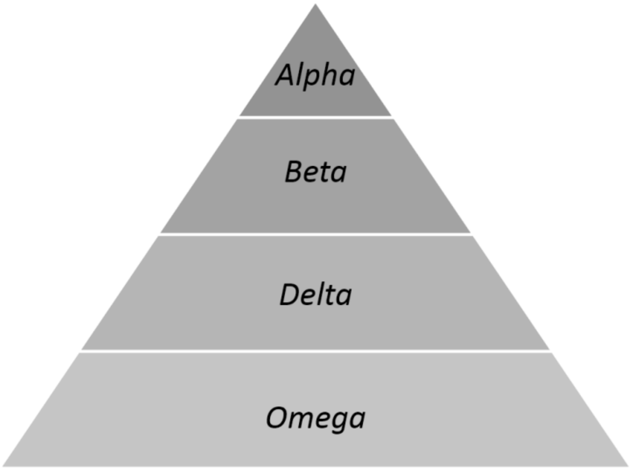

FIGURE 6. Social Hierarchy of grey wolves.

algorithms such as PSO have been used in many proposed protocols [4], [11], [14], and [28]. However, due to some shortcomings in the current optimization algorithms, new optimization algorithms are presented. The grey wolf optimizer (GWO) is one of the recent metaheuristic swarm intelligence methods. It has been widely adopted in many optimization problems due to its impressive characteristics including the following [16]:

1) It needs very few parameters.
2) It is simple, scalable and easy to use.
3) It has a special ability to achieve the right balance between exploration and exploitation.
4) It uses less memory than PSO.

A. GREY WOLF OPTIMIZER OVERVIEW

The inspiration for GWO is the natural behavior and social structure of grey wolves in chasing prey. Each pack of wolves is governed by a hierarchical structure. The most powerful wolf is the alpha, which leads the entire pack. In absence of the alpha wolf, the second most powerful wolf, known as beta wolf, takes the role of the alpha wolf. The delta and omega wolves are the less strong wolves. Fig.6 shows the hierarchical structure of wolf packs.

The grey wolves have a specific intelligent method in chasing and hunting their prey which includes chasing, encircling, harassing, and attacking the prey. In order to model this social structure, the GWO considers the fittest solution the alpha, and second and third best solutions the beta and the delta. The rest of the solutions in the population are considered omega. In this algorithm, the optimization is guided by alpha, beta, and delta wolves while omega wolves follow these three. The algorithm starts by generating a random initial population and repeatedly updates the position of the individuals in the population until a termination criterion is met. Then the best solution which is the alpha wolf is returned as the output of the algorithm. The procedure of updating wolf positions in each iteration involves these steps:

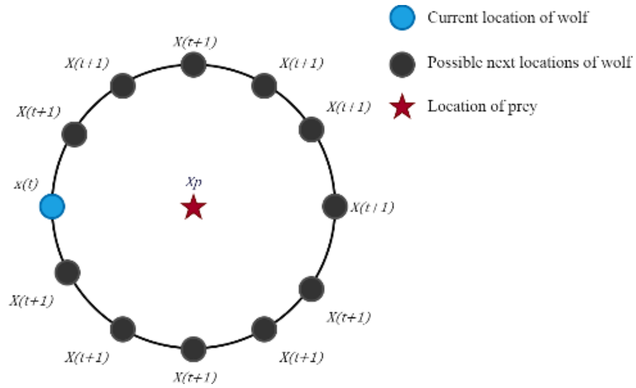

FIGURE 7. How the position of the wolf is updated to allow the wolf to relocate on a circle, sphere or hypersphere around the prey.

1- Encircling the prey:

To mathematically model the encircling of the prey, the GWO employs the following formula:

$$X\,(t+1) = X\,(t) - A.D \tag{9}$$

where $X(t+1)$ is the next location of the wolf and $X(t)$ is its current location. A is a coefficient matrix and D is a vector that depends on the approximated location of the prey which is calculated by the following formula:

$$D = |C.X_p\,(t) - X\,(t)\,| \tag{10}$$

where $C = 2 \times r_2$ and $Xp(t)$ is the current location of the prey and r_2 is a randomly generated vector whose components are in range 0-1. Using these two equations, the wolf will relocate itself on a hypersphere around the prey. The random values are used to simulate different movement speeds of the wolves. To let wolves chase and approach the prey the vector A is defined as:

$$A = 2a \times r_1 - a \tag{11}$$

Components of a are linearly decreased from 2 to 0 over the course of iterations. r_1 is a random vector with components in the range 0-1. The effects of applying these equations to update the position of wolves is that, wolves encircle the prey and change their distance to the prey to achieve both exploration and exploitation. Fig. 7 shows this process.

2- Hunt

To simulate the social structure, omega wolves should follow the alpha, beta and delta wolves. As the position of the prey is not known, it is assumed that alpha, beta and delta wolves, being the three best solutions found, have better knowledge about the location of prey. Therefore the location of prey is approximated by considering the location of alpha, beta and omega wolves and the other wolves are obliged to update their position using the following formula:

$$X\,(t+1) = \frac{X_1 + X_2 + X_3}{3} \tag{12}$$

where

$$\begin{aligned} X_1 &= X_\alpha\,(t) - A_1.D_\alpha \\ X_2 &= X_\beta\,(t) - A_2.D_\beta \\ X_3 &= X_\delta\,(t) - A_3.D_\delta \end{aligned} \tag{13}$$

Initialize the grey wolf population X_i
Initialize a, A and C
calculate the fitness of each individual in population
X_α = the best solution
X_β = the second best solution
X_δ = the third best solution
***while**(t<max number of iterations)*
  ***for** each wolf*
    update the position by equation 12
  end for
  update a, A and C
  calculate the fitness of all wolves
  update X_α, X_β and X_δ
  t = t+1
end while
return X_α

FIGURE 8. **Pseudo code of the gray wolf optimizer algorithm.**

And

$$\begin{aligned} D_\alpha &= |C_1.X_\alpha - X| \\ D_\beta &= |C_2.X_\beta - X| \\ D_\delta &= |C_3.X_\delta - X| \end{aligned} \tag{14}$$

The entire process of the algorithm is given in Fig. 8.

B. CLUSTERING USING GWO

In the proposed protocol, the GWO algorithm is used in order to find a solution that maximizes the fitness function given by (5). To this end, each possible solution for the clustering problem should be presented as an n-dimensional vector. This is because GWO generates a random population of n-dimensional vectors and outputs an n-dimensional vector as the best solution found. This vector is analogous to a particle in PSO and is called a wolf in GWO terminology. In other words, we should define a mapping that maps an n-dimensional vector to a clustering configuration, in order to compute the fitness value of the solution. The number of dimensions of this vector is the number of nodes in the network and each individual dimension i holds the chance of the node i to become CH. A predefined number of nodes for example 5% of nodes with the highest chance are selected as CH and cluster is formed as described in section III-C. This mapping is depicted in Fig.9.

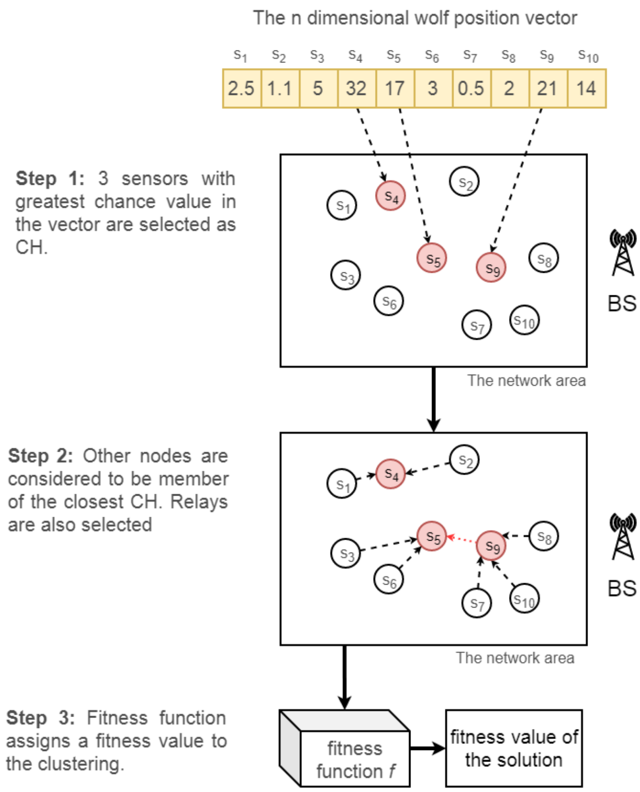

FIGURE 9. **Assume network with 10 nodes and desired 3 CHs. The process to compute the fitness function of the solution given the n-dimensional wolf vector repressing it is shown.**

V. PROTOCOL ANALYSIS

Theorem-1: The best value for relay distance ratio r_0 in order to consume the least amount of energy is $\frac{1}{\sqrt[3]{2}} + 1 \approx 1.793$.

Proof: by considering the model shown in Fig.3, the amount of energy that a CH consumes if it uses a relay is given by (15).

$$E^{ch}_{dual-hop} = E_{Tx}(L, \frac{D}{r_0}) \tag{15}$$

That is, the energy consumed for transmitting an L bit message over distance D/r_0. The energy that the relay consumes is given by (16).

$$E^{relay}_{dual-hop} = 2 \times E_{Tx}(L, (D - \frac{D}{r_0})) \tag{16}$$

This corresponds to sending 2 messages of length L from relay to the BS.

Therefore the total energy consumed by a CH and a relay in dual-hop communications is:

$$E^{total}_{dual-hop} = E_{TX}\left(L, \frac{D}{r_0}\right) + 2 \times E_{Tx}(L, (D - \frac{D}{r_0})) \tag{17}$$

To achieve the best energy saving, the expression given by (17) should be minimized. By considering multipath fading model and expanding functions from (3) we have:

$$E^{total}_{dual-hop} = L \times (E_{mp} \times \left(\left(\frac{D}{r_0}\right)^4 + 2 \times \left(D - \frac{D}{r_0}\right)^4\right) + E_{elec}) \tag{18}$$

This can be written as:

$$E^{total}_{dual-hop} = L \times (E_{mp} \times \left(\frac{D^4}{r_0^4} + 2 \times (\frac{D^4(r_0-1)^4}{r_0^4})\right) + E_{elec}) \tag{19}$$

Because L, E_{mp}, and E_{elec} are constant, (19) is minimized where h_1 which is given by (20) is minimized.

$$h_1 = \frac{D^4}{r_0^4} + 2 \times (\frac{D^4(r_0 - 1)^4}{r_0^4}) \tag{20}$$

Furthermore, because D is also a non-zero constant, to minimize h_1 we can minimize h_2 which is given by (21).

$$h_2 = \frac{h_1}{D^4} = \frac{1 + 2 \times (r_0 - 1)^4}{r_0^4} \tag{21}$$

The minimum value of h_2 is 0.173 at $r_0 = \frac{1}{\sqrt[3]{2}} + 1 \approx 1.793$. Therefore, the best position for a relay in order to save most energy is found.

Theorem-2: The best value for relay distance ratio r_0 in order to make CH and its relay consume an equal amount of energy is approximately 1.8408.

Proof: Similar to previous proof, to achieve equal value for energy consumed by relay and CH the difference between the two should be minimized.

$$minimize\ g = |E^{ch}_{dual-hop} - E^{relay}_{dual-hop}| \tag{22}$$

By replacing values from (3) we have

$$g = |L \times (E_{mp} \times \left(\frac{D^4}{r_0^4} - 2 \times (D - \frac{D}{r_0})^4\right) - E_{elec})| \tag{23}$$

Because L, E_{mp}, D, and E_{elec} are none-zero constants, g is minimized where g_0 is zero. g_0 is given by (24).

$$g_0(r_0) = \frac{1 - 2 \times (r_0 - 1)^4}{r_0^4} \tag{24}$$

The root of this function is the best value for r_0 to ensure balanced energy consumption among relay and CH. One suitable root for this function is 1.8408, therefore choosing $r_0 = 1.8408$ results in equal energy consumption for CH and its relay.

To choose the best r_0 overall, we used the average of these two numbers to achieve both goals simultaneously. Therefore our final value for r_0 is $(1.8408+1.793)/2 = 1.8169$.

Theorem-3: The complexity of control messages in the protocol is $O(N)$.

Proof: We assume the number of nodes in the network is N. In each round of the protocol N Node-MSG messages are sent to inform the BS of the state of the nodes, each node also sends a Join-MSG to the selected CH. In addition to these, each CH sends a CH-ADV message as well as a either a Relay-ADV or a RJoin-MSG to form dual hop routing. If we consider that 5% of nods are CHs then there are *N/20* CHs in the network, total control packet required to form clustering is $2N + 2 * (\frac{\mathrm{N}}{20}) = \frac{21}{10}N$. Therefore the overall complexity of control messages of the network is $O(N)$.

TABLE 1. **Common parameters used in simulations.**

Type	Parameter	Value
Network	Initial energy	0.5J
	Percentage of CHs	5%
	Packet length	4000bits
	Message length	200bits
Energy model	E_{elec}	50nJ/bit
	E_{fs}	10pJ/bit/m^2
	E_{mp}	0.0013pJ/bit/m^4
	E_{DA}	5nJ/bit/signal
	d_0	87m
GWO	GWO population size	30
	GWO iterations	50

TABLE 2. **Parameters of scenarios.**

	Area	BS location	Nodes
Scenario 1	200m*200m	(0,0)	100
Scenario 2	200m*200m	(100,100)	150
Scenario 3	100m*100m	(150,100)	100
Scenario 4	300m*300m	(150,500)	500

VI. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we present the results of simulating our approach against LEACH-C, PSO-ECHS, CHIPIPSO, PSO-UCF, UCRA-GOS and SMOTECP protocols. We employed C++ to write our simulation codes. The metrics used to evaluate performance of the protocols are network lifetime and energy consumption. Network lifetime is also represented as the round at which the first node dies (FND), the round at which half of the nodes are dead (HND) and the last round (LND). TABLE 1 shows common parameters used in simulation.

The number of nodes, network area size, and location of the BS are the three main parameters that affect the lifetime of network [11]. The protocols have been simulated in 4 different scenarios shown in TABLE 2.

In the first scenario, a WSN with 100 nodes and parameters presented in TABLE 1 and 2 is simulated. Network lifetime metrics we obtained from this simulation is shown in Fig.10.

Our protocol has managed to improve FND, HND and LND against all compared protocols. Our proposed protocol does not execute cluster setup phase in every round which saves the energy that would be required to form clustering due to the exchange of control packets. PSO-UCF has a

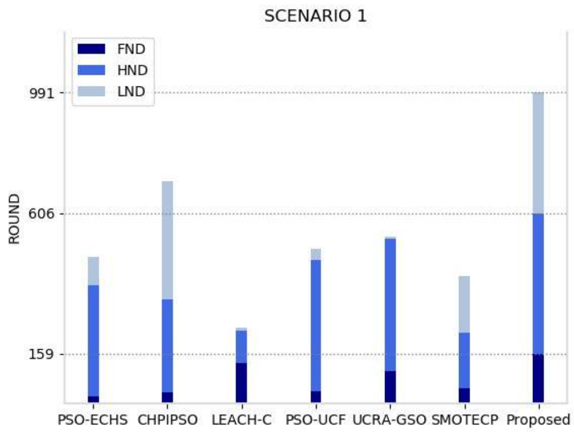

FIGURE 10. Lifetime metrics of scenario 1.

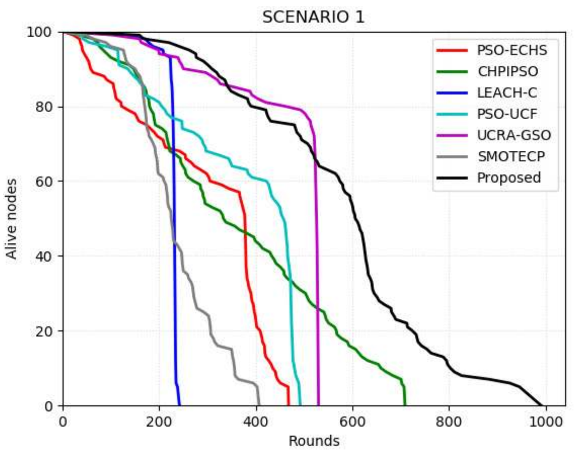

FIGURE 11. Alive nodes per round for scenario 1.

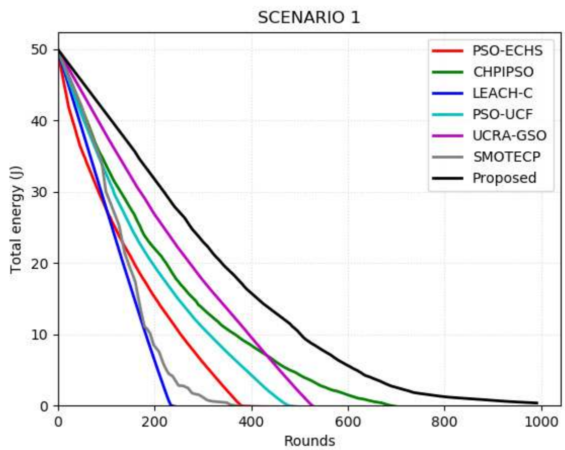

FIGURE 12. Total energy for scenario 1.

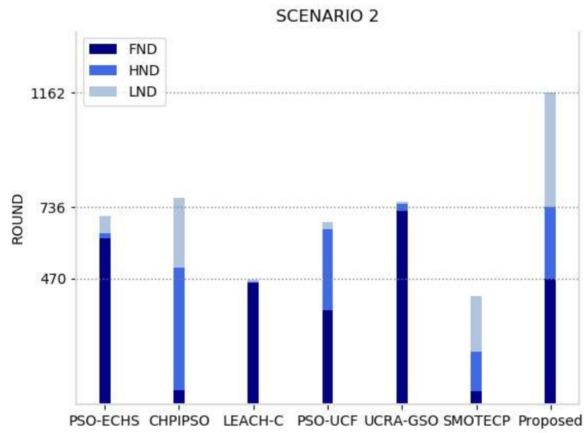

FIGURE 13. Lifetime metrics of scenario 2.

similar feature which manifests in the relatively competitive results. Another feature that leads to a better network lifetime in our protocol is dual-hop routing and carefully selecting relays. The CHIPIPSO protocol also uses dual-hop routing and selects relays using PSO which results in extended network lifetime compared to other protocols.

Another examined metric is the exact number of alive nodes per round which is plotted in Fig. 11.

The proposed protocol manages to keep network energy efficient even after a significant number of nodes are dead but other protocols do not manage this situation well. This is because the proposed protocol balances energy among far and near CH very well using the presented relay selection method.

The last metric that is considered is the total remaining energy of nodes which is plotted for scenario 1 in Fig. 12. Because the proposed protocol approximates the actual energy that would be consumed under any certain clustering and uses this to rate solutions, it finds better solutions in terms of energy consumption.

The simulation parameters for second, third and fourth scenarios are listed in TABLE 2. In Scenario 2 node density increases and BS is placed in the center of the field. This minimizes the effect of hot spot problem [4]. Fig. 13 shows the lifetime metrics for scenario 2. Several protocols have shown improvement over scenario 1 in FND metric because of the aforementioned reasons. In fact, PSO-ECHS and UCRA-GSO perform better than the proposed protocol in FND metric but both of these protocols fail to keep nodes alive much longer. The alive nodes per round plot for scenario 2 is shown in Fig.14.

Fig. 15 shows the lifetime metrics for scenario 3. Compared to scenario 1, network density is increased and the BS is moved further outside the field. The proposed protocol outperformed other protocols in all metrics except UCRA-GSO protocol which preformed 4% better in FND. The protocols which use dual-hop routing show better results in this scenario because BS is located further. Fig. 16 shows alive nodes per round plot for scenario 3.

In the last scenario, the number of nodes is increased to 500 while the network area became larger and BS is located further outside the field. In this scenario scalability of the protocols is tested. The proposed protocol performed better than compared protocols in all three lifetime metrics. Fig. 17 shows network lifetime metrics for scenario 4.

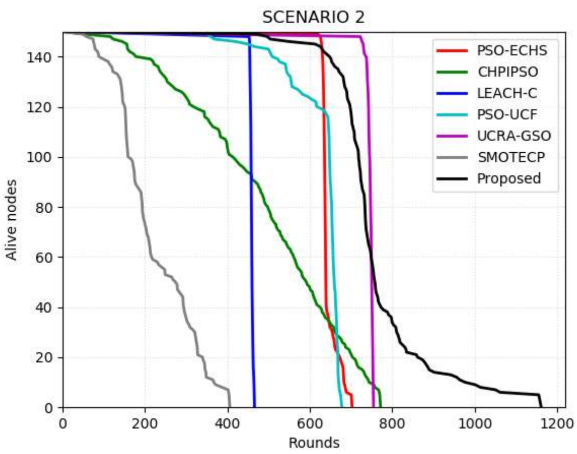

FIGURE 14. Alive nodes per round for scenario 2.

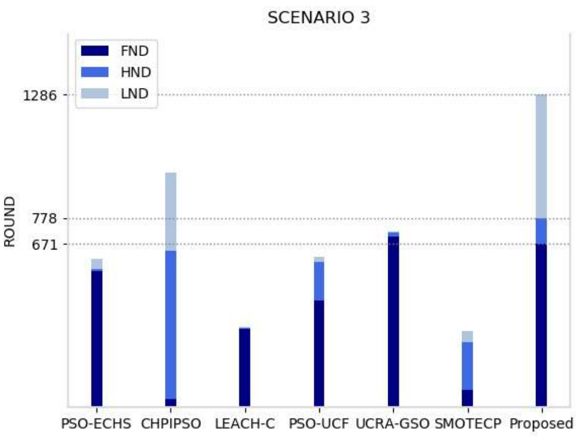

FIGURE 15. Lifetime metrics of scenario 3.

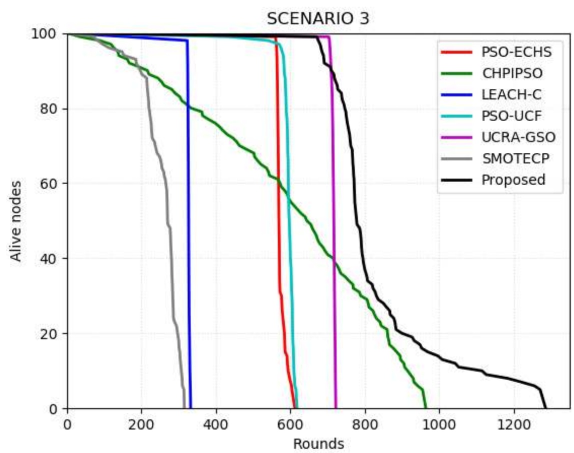

FIGURE 16. Alive nodes per round for scenario 3.

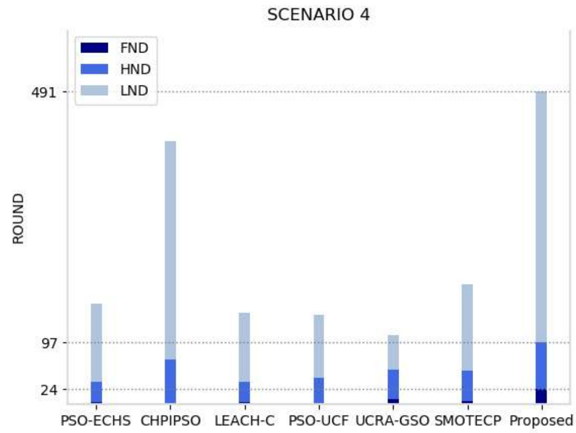

FIGURE 17. Lifetime parameters of scenario 4.

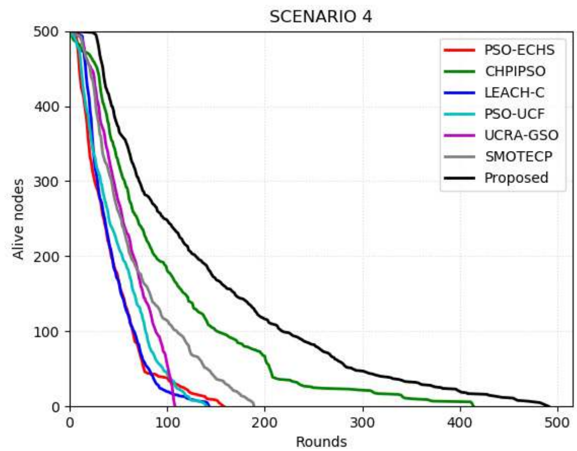

FIGURE 18. Alive nodes per round for scenario 4.

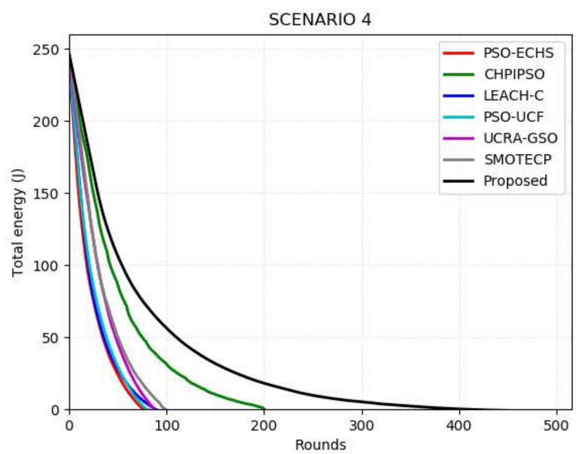

FIGURE 19. Total energy for scenario 4.

The alive nodes per round plot are also presented in Fig. 18. Higher node density in this scenario results in a more gradual death of nodes across all the protocols but because the BS is located further, the energy consumption is also higher. The total energy per round plot for scenario 4 is shown in Fig. 19.

TABLE 3 presents the FND, HND and LND for the proposed, PSO-ECHS, CHIPIPSO, LEACH-C, PSO-UCF, UCRA-GSO and SMOTECP protocols in all scenarios. The improvement achieved by the proposed protocol against the best of compared protocols is also noted. The proposed

TABLE 3. Comparison of lifetime parameters.

Scenario	Protocol	FND	HND	LND
Scenario 1	PSO-ECHS	22	378	468
	CHIPIPSO	35	333	709
	LEACH-C	130	232	242
	PSO-UCF	39	457	492
	UCRA-GSO	105	526	530
	SMOTECP	49	255	407
	PROPOSED	159	606	991
	IMPROVED	**22%**	**15%**	**39%**
Scenario 2	PSO-ECHS	620	638	703
	CHIPIPSO	52	509	772
	LEACH-C	453	459	466
	PSO-UCF	351	654	678
	UCRA-GSO	723	749	755
	SMOTECP	49	197	406
	PROPOSED	470	736	1162
	IMPROVED	**-34%**	**-1%**	**50%**
Scenario 3	PSO-ECHS	560	570	612
	CHIPIPSO	35	644	964
	LEACH-C	323	327	332
	PSO-UCF	439	597	618
	UCRA-GSO	703	717	722
	SMOTECP	71	270	315
	PROPOSED	671	778	1286
	IMPROVED	**-4%**	**8%**	**66%**
Scenario 4	PSO-ECHS	4	35	158
	CHIPIPSO	1	70	414
	LEACH-C	4	36	143
	PSO-UCF	1	41	141
	UCRA-GSO	8	55	108
	SMOTECP	5	53	189
	PROPOSED	24	97	491
	IMPROVED	**200%**	**38%**	**18%**

protocol improved FND in all scenarios by 18% to 66%. It also improved HND except against UCRA-GSO in scenario 2, in other cases it improved HND by 8% to 38%. The proposed protocol does not always show improvement in terms of FND. It is because the protocol tries to use a node as CH for several rounds which results in early FND.

VII. CONCLUSION

In this paper, we presented a new clustering protocol that uses the grey wolf optimizer and a single term fitness function. The fitness function is a prediction of the actual energy consumption in the network under each considered solution. The proposed protocol also prevents waste of energy due to unnecessary execution of the cluster setup phase in rounds where the current clustering is still good enough. To prevent fast energy depletion of the distant nodes from the BS, the relay selection is proved to minimize total energy consumption while balancing energy consumption between CHs and relays. It has been shown that the proposed protocol can yield competitive results in comparison to a few of similar recent protocols. However, the proposed protocol may not be suitable for application where FND has a significant effect on performance of the network. It is also not suitable in applications where fault is critical because no fault tolerance mechanism is added to the protocol. Therefore future works can focus on adding fault tolerance to the protocol. And because no quality of service (QoS) metric other than lifetime is considered, future research should focus on adapting the protocol with QoS considerations.

REFERENCES

[1] D. C. Hoang, P. Yadav, R. Kumar, and S. K. Panda, "Real-time implementation of a harmony search algorithm-based clustering protocol for energy-efficient wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 774–783, Feb. 2014.

[2] W. Xiang, N. Wang, and Y. Zhou, "An energy-efficient routing algorithm for software-defined wireless sensor networks," *IEEE Sensors J.*, vol. 16, no. 20, pp. 7393–7400, Oct. 2016.

[3] J. Wang, X. Gu, W. Liu, A. K. Sangaiah, and H.-J. Kim, "An empower hamilton loop based data collection algorithm with mobile agent for WSNs," *Hum.-Centric Comput. Inf. Sci.*, vol. 9, no. 1, p. 18, Dec. 2019.

[4] T. Kaur and D. Kumar, "Particle swarm optimization-based unequal and fault tolerant clustering protocol for wireless sensor networks," *IEEE Sensors J.*, vol. 18, no. 11, pp. 4614–4622, Jun. 2018.

[5] D. R. Edla, M. C. Kongara, and R. Cheruku, "SCE-PSO based clustering approach for load balancing of gateways in wireless sensor networks," *Wireless Netw.*, vol. 25, no. 3, pp. 1067–1081, Apr. 2019.

[6] J. Wang, Y. Gao, X. Yin, F. Li, and H.-J. Kim, "An enhanced PEGASIS algorithm with mobile sink support for wireless sensor networks," *Wireless Commun. Mobile Comput.*, vol. 2018, Nov. 2018, Art. no. 9472075.

[7] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, and L. Hanzo, "Network-lifetime maximization of wireless sensor networks," *IEEE Access*, vol. 3, no. 2015, pp. 2191–2226, Nov. 2015.

[8] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, and L. H. Hanzo, "A survey of network lifetime maximization techniques in wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 828–854, 2nd Quart., 2017.

[9] M. S. Bahbahani and E. A. Alsusa, "A cooperative clustering protocol with duty cycling for energy harvesting enabled wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 101–111, Jan. 2018.

[10] J. Wang, Y. Gao, W. Liu, A. K. Sangaiah, and H.-J. Kim, "An intelligent data gathering schema with data fusion supported for mobile sink in wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 15, no. 3, Mar. 2019, Art. no. 1550147719839581.

[11] Y. Zhou, N. Wang, and W. Xiang, "Clustering hierarchy protocol in wireless sensor networks using an improved PSO algorithm," *IEEE Access*, vol. 5, pp. 2241–2253, 2017.

[12] 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.

[13] P. S. Mann and S. Singh, "Improved metaheuristic based energy-efficient clustering protocol for wireless sensor networks," *Eng. Appl. Artif. Intell.*, vol. 57, pp. 142–152, Jan. 2017.

[14] P. C. Rao, P. K. Jana, and H. Banka, "A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks," *Wireless Netw.*, vol. 23, no. 7, pp. 2005–2020, Oct. 2017.

[15] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014.

[16] H. Faris, I. Aljarah, M. A. Al-Betar, and S. Mirjalili, "Grey wolf optimizer: A review of recent variants and applications," *Neural Comput. Appl.*, vol. 30, no. 2, pp. 413–435, Jul. 2018.

[17] A. S. M. Hosen and G. H. Cho, "An energy centric cluster-based routing protocol for wireless sensor networks," *Sensors*, vol. 18, no. 5, p. 1520, May 2018.

[18] 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.*, Jan. 2000, p. 10.

[19] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, Oct. 2002.

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

[21] N. Mittal, U. Singh, R. Salgotra, and B. S. Sohi, "A Boolean spider monkey optimization based energy efficient clustering approach for WSNs," *Wireless Netw.*, vol. 24, no. 6, pp. 2093–2109, Aug. 2018.

[22] J. C. Bansal, H. Sharma, S. S. Jadon, and M. Clerc, "Spider monkey optimization algorithm for numerical optimization," *Memetic Comput.*, vol. 6, no. 1, pp. 31–47, Mar. 2014.

[23] J. Wang, Y. Gao, W. Liu, A. K. Sangaiah, and H.-J. Kim, "An improved routing schema with special clustering using PSO algorithm for heterogeneous wireless sensor network," *Sensors*, vol. 19, no. 3, p. 671, 2019.

[24] Y. Xiuwu, L. Qin, L. Yong, H. Mufang, Z. Ke, and X. Renrong, "Uneven clustering routing algorithm based on glowworm swarm optimization," *Ad Hoc Netw.*, vol. 93, Oct. 2019, Art. no. 101923.

[25] N. A. Al-Aboody and H. S. Al-Raweshidy, "Grey wolf optimization-based energy-efficient routing protocol for heterogeneous wireless sensor networks," in *Proc. 4th Int. Symp. Comput. Bus. Intell. (ISCBI)*, Sep. 2016, pp. 101–107.

[26] J. Wang, Y. Cao, B. Li, H.-J. Kim, and S. Lee, "Particle swarm optimization based clustering algorithm with mobile sink for WSNs," *Future Gener. Comput. Syst.*, vol. 76, pp. 452–457, Nov. 2017.

[27] J. Wang, Y. Gao, W. Liu, W. Wu, and S.-J. Lim, "An asynchronous clustering and mobile data gathering schema based on timer mechanism in wireless sensor networks," *Comput., Mater. Continua*, vol. 58, no. 3, pp. 711–725, 2019.

[28] D. Chandirasekaran and T. Jayabarathi, "Cat swarm algorithm in wireless sensor networks for optimized cluster head selection: A real time approach," *Cluster Comput.*, vol. 22, pp. 11351–11361, Sep. 2019.

S. M. MAHDI H. DANESHVAR received the B.S. and M.S. degrees in computer engineering from Imam Reza University, Iran, in 2015 and 2019, respectively. His research interests include wireless sensor networks, soft computing, computer networks, and software engineering.

PARDIS ALIKHAH AHARI MOHAJER received the B.S. and M.S. degrees in biomedical engineering and communication engineering from Imam Reza University, Iran, in 2017 and 2019, respectively. Her research interests include wireless sensor networks, biosensor, and intelligent systems.

SAYYED MAJID MAZINANI was born in Mashhad, Iran, in 1971. He received the bachelor's degree in electronics from Ferdowsi University, Mashhad, Iran, in 1994, the master's degree in remote sensing and image processing from Tarbiat Modarres University, Tehran, Iran, in 1997, and the Ph.D. degree in wireless sensor networks from Ferdowsi University, in 2009. He worked at IRIB, from 1999 to 2004. He was the Head of the Department of Electrical and Computer Engineering, Imam Reza International University, Mashhad, from 2009 to 2012, where he is currently an Assistant Professor with the Faculty of Engineering. His research interests include computer networks, wireless sensor networks, and smart grids.