Energy Efficient Multi-Clustering Using Grey Wolf Optimizer in Wireless Sensor Network

Maryam Ghorbanvirdi

Department of Electrical Engineering
Imam Reza International University
Mashhad, Iran
maryam.ghorbanvirdi@imamreza.ac.ir

Sayyed Majid Mazinani[*]

Department of Electrical Engineering
Imam Reza International University
Mashhad, Iran
smajidmazinani@imamreza.ac.ir

Abstract—**The most important challenge in wireless sensor networks is to extend the network lifetime, which is directly related to the energy consumption. Clustering is one of the well-known energy-saving solutions in WSNs. To put this in perspective, the most studies repeated cluster head selection methods for clustering in each round, which increases the number of sent and received messages. what's more, inappropriate cluster head selection and unbalanced clusters have increased energy dissipation. To create balanced clusters and reduce energy consumption, we used a centralized network and relay nodes, respectively. Besides, we applied a metaheuristic algorithm to select the optimal cluster heads because classical methods are easily trapped in local minimum. In this paper, the Grey Wolf Optimizer(GWO), which is a simple and flexible algorithm that is capable of balancing the two phases of exploration and exploitation is used. To prolong the network lifetime and reduce energy consumption in cluster head nodes, we proposed a centralized multiple clustering based on GWO that uses both energy and distance in cluster head selection. This research is compared with classical and metaheuristic algorithms in three scenarios based on the criteria of "Network Lifetime", "Number of dead nodes in each round" and "Total Remaining Energy(TRE) in the cluster head and relay nodes. The simulation results show that our research performs better than other methods. In addition, to analyze the scalability, it has been evaluated in terms of "number of nodes", "network dimensions" and "BS location". Regarding to the results, by rising 2 and 5 times of these conditions, the network performance is increased by 1.5 and 2 times, respectively.**

Keywords: Multi-clustering; Centralized; Energy efficient; Grey wolf algorithm; Wireless sensor network.

I. INTRODUCTION

A wireless sensor network is a set of independent devices that measure and monitor environmental conditions using sensors. Since, these sensors need to stay in the network for a long time and also charging or replacing them causes a problem, energy conservation has become one of the most important challenges in these networks. Clustering is the most popular energy-saving solutions in WSN [1], [2], and it's called a process in which the samples are divided into groups whose members are similar. There are many criteria for similarity, such as distance-based clustering. Clustering is used in many fields of science and engineering. Due to the wide range of usage,

---

* Corresponding Author

researchers are trying to design new clustering algorithms and improve the performance of existing algorithms using new methods[3]. Each of these algorithms has its advantages. A common criterion for all of these algorithms is the stability of the algorithm throughout the lifetime of the network. Most studies repeated cluster head selection methods in each round, which increases the number of sent and received messages. This, reduces the lifetime of the network by increasing the energy consumption of the nodes[4]. Classical and metaheuristic methods are two broad areas of clustering algorithms. Classical clustering is divided into five types: area-based, hierarchical, density-based, grid-based, and model-based. One disadvantage of these methods is to get stuck in local optimum easily. In the last few decades, many metaheuristic methods have been used to overcome this weakness. Metaheuristic methods can provide near-optimum solutions in less time than classical methods[3], [5], [6]. Metaheuristic optimizers have been known as effective ways of solving complex optimization problems. Over the past two decades, optimization techniques such as the Genetic Algorithm, Ant Colony Algorithm, and Particle Swarm Optimization have attracted not only computer scientists, but also scientists in other fields[3], [6], [7]. The question that may arise is why metaheuristics have become so popular? The answer to this question can be divided into four main reasons: simplicity, flexibility, freedom of derivative mechanism and avoidance of trapping in local optimum. Metaheuristic methods are Simple and mainly inspired by physical phenomena, animal behaviors, or evolutionary concepts. The flexibility of metaheuristics demonstrates the use of these methods in various ways without making any specific changes to the algorithm structure[7], [8]. Since problems are considered as black boxes in the metaheuristic method, they are easily applied to different problems. In other words, only the inputs and outputs of a system are important for a metaheuristic, so the only important issue for the designer is, how to apply the problem to the method. In fact, in metaheuristic algorithms, optimization done by using a set of solutions (population) [7]. Regardless of the difference between the metaheuristic algorithms, one common feature among them is that the search process is divided into two distinct phases: Exploration and Exploitation. The Exploration phase is the process of discovering the range of possible answers. Adding pseudo-random operators to such algorithms allows global search across the entire search space. While Exploitation, local search capabilities in areas that have been obtained in the exploration phase. However, because of the pseudo-random nature of metaheuristic algorithms, achieving the proper balance between these two phases has become a challenge in such protocols[5]. In this study, the grey wolf algorithm proposed by Mirjalili et al. [9] is used to find optimum solutions to cost functions. The grey wolf has few parameters. The algorithm is simple, easy to use, flexible and scalable, and has the special ability to strike the right balance between the two phases of Exploration and Exploitation[9]. GWO Like any other

metaheuristic starts with generating a random population of solutions. What distinguishes this algorithm from other algorithms is that this algorithm, with two parameters, applies a robust control function to avoid local optimum. Also, the mathematical model of this algorithm is new in its kind and allows solutions to be searched in a n-dimensional space. GWO has a position vector, so this algorithm requires less memory than an algorithm such PSO (position and velocity), besides, the algorithm keeps the best three solutions in each round of the network while the PSO keeps only one solution[10]. Another advantage of this algorithm is the capability to be developed and applied in most fields[10]–[12]. The rest of the paper is organized as following sections: in Sec. II. the related works were done, in addition, the classical and the metaheuristic protocols are organized in this section. The system model is described in Sec. II. In Sec. IV. the protocol description is presented. The simulation and analysis are organized in Sec. V. and finally, Sec.VI. deals with the results.

II. RELAYED WORK

The development of wireless sensor networks has led to a large body of research; In such variable conditions, designing and developing efficient protocols for research on wireless sensor networks is a challenge. Besides, classical and metaheuristic approaches provide solutions for wireless sensor networks resource constraints [13]. We introduce some of these methods:

A. Classical Protocols

To date, a great deal of research has been done to improve energy consumption using protocols based on clustering algorithms. Probabilistic clustering algorithms, over higher energy efficiency, usually speed up execution or convergence times and reduce the volume of exchanged messages [14]. Here, a protocol has played a key role in the emergence of many new algorithms. In 2000, Heinzelman et al. [15] proposed an algorithm called LEACH(Low-energy adaptive clustering hierarchy), which took a special place among routing protocols in sensor networks and has done a lot of optimization of distributed solutions so far [13]. LEACH is a classic example of data routing, the function of LEACH is organized in rounds, each round consisting of a setup phase and a data transmission phase. During the setup phase, the nodes organize themselves into clusters with a node as the cluster head. In the data transmission phase, the cluster head collects the data and sends it to the base station after the data is merged. The information transfer phase time is much longer than the configuration phase to compensate for the overhead due to cluster formation. Therefore LEACH has presented a good model in which local algorithms and data collection can randomly select their cluster heads, which reduces the overhead of information, and provides a reliable set of data to the end-user, since this distributed protocol did not form balanced clusters, so its central model was presented [16]. Manjeshwar et al. [17] in 2001

introduced a hierarchical protocol called TEEN(Threshold sensitive Energy Efficient sensor Network protocol), designed for situations where sudden changes in the measured parameter are to be reported and for appropriate applications requiring periodic information reporting and it is used only for specific networks and conditions. In 2002, Heinzelman et al. [18] introduced the centralized leach algorithm, called LEACH-C. The steady-state phase is similar to leach, but the setup phase is completely different, in which each sensor node sends its energy and location information to the base station and the base station calculates the average energy level of the network nodes. Cluster head selection criteria is based on energy level greater than the average. Therefore, the choice of a cluster head plays a vital role in the lifetime of the sensor network, early discharge energy of the sensors being the weaknesses of this algorithm. In 2002 Lindsay et al. [19] introduced a chain-based data collection protocol called PEGASIS (Power-Efficient Gathering in Sensor Information Systems). In this protocol, instead of forming different clusters, a chain of interconnection is established between all the sensors in the network. The weaknesses of this protocol are too much delay. In 2004, Younes et al. [20] proposed a protocol called HEED (A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks) that periodically selects cluster heads based on two parameters: residual energy and Proximity to the neighbors or node degree. Although high latency is one of the disadvantages of this work, this protocol is effective in increasing the network lifetime and displaying scalable data. In 2019, Al-Hamidi et al. [21] proposed a centralized routing method called the EACCC, which consists of two steps: In the first step, all nodes send their remaining energy and location information to the base station, and the second step involves the three phases of selecting the cluster head, forming the cluster and sending the information. This method develops central clustering techniques to achieve greater energy, longer network life, and greater flexibility. They do this by reducing the number of sent and received messages, and increasing the number of live nodes. In 2019, Zhong et al. [22] introduced an improved algorithm based on leach called LAN(Load Balancing Network). The algorithm is like leach periodically, and the network function consists of two phases. To reduce resource overhead, the steady-state phase duration is longer than the set-up phase, which causes overloads in the setup phase. The residual energy and the distribution of the cluster heads play a role in determining the cluster heads. In this protocol, the optimum number of cluster heads of the current network is estimated and accordingly, the network is divided into square sections. This protocol improves network features over a low coverage area compared to leach but is not effective for networks with large coverage levels. In 2019, Azizi et al. [23] have designed a new method for calculating energy-efficient routing. This method is inspired by LEACH. In this method, a threshold level is used to calculate the best possible candidates of cluster heads and then determine the cluster heads candidates using the energy and distance criteria to the base station. In this way, the stronger cluster head plays the role of a node gateway between the cluster heads and the base station. The method for both homogeneous and heterogeneous networks was tested and both have shown good performance.

B. Metaheuristics Protocols

The main problem of the classical methods is getting stuck in local optimum. Therefore, researchers have used metaheuristic algorithms to overcome the limitations of classical algorithms [3]. In the field of metaheuristic, many algorithms are mainly inspired by nature [5], [24]. Some of the most popular are: SA [25], ACO [26], PSO [27], GA [28], GWO [9].

In 1991, Salim et al. [25] proposed an algorithm called simulated annealing for clustering, where they theoretically proved that a global solution to the clustering problem could be obtained. The main weakness of this method is the parameter setting. In 1992, Holland [28] introduced a new algorithm called genetic algorithm for clustering problems. The Genetic algorithm is an optimization technique and it tries to find the values of the input that will produce the best output. Of course, it works well if the response space is consistent. This algorithm is iterative and includes five sections: initial population production, fitness function, selection, crossover, and mutations. In 1995, Eberhart et al. [27] proposed a particle swarm optimization(PSO) method based on the social behavior of birds. Unlike simple mathematics, it has been used in many areas of different optimization problems. One of the strength of this algorithm is better convergence and data transfer rates, while it is ineffective for data sets that overlap and suffer from network overload [7], [13], [29]. In 1998, Shi et al. [30] added the inertia weight parameter to the initial version of the particle swarm and developed its model. The simulation results show the positive effect of this parameter on PSO performance. It has given a great chance of finding the global optimum. In 1999, Dorigo et al. [26] proposed a metaheuristic approach that simulates the ant's behavior to find the shortest path from the nest to the food source, called the ACO (Ant Colony Optimization). This method shows good convergence, but suffers from low coverage. In 2004, Shlovar et al. [31] compared ant colony performance with GA, TS, SA algorithms, and the results show better simulation performance of this algorithm. Caraboga et al. [32] proposed an algorithm based on honey bee searching behavior in 2005, It was first proposed to solve nonlinear optimization problems using only the usual control parameters such as colony size and the maximum number of cycles. This algorithm tries to balance the two phases of exploration and exploitation, by combining local search and global search. In 2007, Karaboga et al. [36] developed this algorithm to solve definitive optimization problems and compared it with PSO and DE. The results, had better performance in cluster quality and processing time. Despite the applications of this method, the parameter setting is the main weakness of this method.

Nayak [33] in 2017 uses GA algorithm to generate optimal cluster centers compared to the classical clustering types. In 2018, Reddy et al. [34] have designed a PSO-based routing method that proposed a new cost function based on relay nodes, cluster head distance to the base station and network load factor. In 2019, Arora et al. [35], using the Ant Colony called AOSTEB (ACO optimized self-organized tree-based energy balance), developed a self-organized routing method that selects its path based on energy and the shortest distance. The protocol consists of three phases: clustering, multi-path creation, and sending information. During the formation of clusters, the optimum number of nodes as cluster heads are selected. Next, multi-path between the cluster head and the members of the cluster are followed by the ant colony, and eventually, an optimal dynamic path is created based on energy and the minimum distance to send information.

TABLE I. COMPARISON OF SOME PROTOCOLS [36]

Protocol	Classification	Energy Efficiency	Advantages	Disadvantages
LEACH [38]	*classical*	*average*	*self-organized*	*high cost communications*
LEACH-C [18]	*classical*	*average*	*optimized cluster formation*	*upcoming offload energy*
PEGASIS [19]	*classical*	*average*	*decrease communications*	*high delay*
HEED [20]	*classical*	*average*	*low cost communications*	*high delay*
TEEN [17]	*classical*	*average*	*sending changes quickie*	*special usage*
PSO [34]	*metaheuristic*	*good*	*high data transmit rate*	*network overload*
AOSTEB [30]	*metaheuristic*	*good*	*high convergence*	*low coverage*
GWO [9]	*metaheuristic*	*Very good*	*few parameters, flexible, simple, right balance in exploration and exploitation*	*appropriate for large networks*
SA [25]	*metaheuristic*	*good*	*global optimum*	*parameter setting*
ABC [32]	*metaheuristic*	*good*	*low cost process*	*parameter setting*

It is clear that classical approaches, as well as metaheuristic approaches, are able to maintain network lifetime under some circumstances, but for two major reasons there is no guarantee that the selected node as the cluster head is the best choice: First, some nodes with the lowest energy are likely to be selected as the cluster head, which in turn increases energy consumption and Second, some nodes are inappropriate because of their location, for example, if the node at the boundary of a network is selected as the cluster head, the energy dissipation increases due to the distance from the base station[7]. Classical approaches work well in self-organization, load balancing with minimal overload, but moderate in energy efficiency, while metaheuristic algorithms are best in energy efficiency with long network lifetime. Therefore, metaheuristic approaches for energy-efficient solutions in WSNs should be further explored and improved [5], [6], [36].

In this paper, we aim to fix the weaknesses of previous studies and try to further reduce the disadvantages of cluster head selection by utilizing the relay nodes as well as the grey wolf optimization algorithm. Using the relay node reduces energy consumption on cluster heads. In fact, by assigning a relay node to each cluster head, the cluster heads no longer need to consume extra energy for their next step[7]. The choice of cluster heads and relay nodes is based on cost functions through energy and distance criteria that can be formulated as an NP-hard problem and the GWO is considered to achieve the optimum solution. In Table I you will briefly see some of these protocols and their advantages and disadvantages[3], [36].

III. MODELS FOR NETWORK AND RADIO ENERGY

A. Network Model

In a wireless sensor network, the sensor nodes need to be stay in the network for a long time. They collect data from environmental conditions, communicate with the base station and other nodes. They lose some of their energy per message, which causes them died prematurely. Also, charging or replacing them has some issues. So Energy saving has become a big challenge in wireless sensor networks. The energy consumption of the nodes can be significantly reduced by clustering, which increases network lifetime. Therefore, it would be useful to propose a protocol that would reduce the energy dissipation of the network as much as possible. Before examining the details of the proposed method, we examine the hypotheses of the system model. Which include:

- All nodes are homogeneous.
- All nodes have equal initial energy.
- Nodes have uniform random distributions.
- The number of clusters is 5% of the total number of nodes.
- Each node in the cluster has its relay node.
- After the distribution, all nodes and base station remains motionless.
- Selection of cluster heads and relay nodes is performed in each round.
- The network is centralized.
- The nodes are aware of their position as well as the position of the other nodes and the base station.
- Euclidean distance is used to determine all distances.
- The cluster heads collect the data from the nodes and then send them to their special relay

node and eventually move from there to the base station.

- The network using a TDMA schedule.
- Each node is marked with an index based on its position.

B. Radio Energy Model

Since most energy is lost during transmission, an energy optimization method is used. The energy model [4], considered in this algorithm is illustrated in Fig.1. The transmitter consumes energy for radio electronics and amplifier components, While the receiver consumes energy only in radio electronics. The amount of energy required to transmit l-bits (from the transmitter to the receiver over a distance d) is shown in (1), (2) and (3). Where E_{Tx} represents the transmit energy and E_{elec} represents the energy released per bit to execute the transmitter or receiver circuit. ε_{fs} and ε_{emp} are transmitter amplifier features. Specifically, ε_{fs} is used for free space and ε_{emp} is used for multipath. When the distance between the transmitter and receiver is less than the threshold value d_0 the free space model d^2 is used. Otherwise, the multi-path channel model d^4 is used. E_{Rx} Indicates the amount of power consumed to receive l-bits of data, refer to in (4). Refer to (5) the threshold value of d_0 which is the ratio of ε_{fs} to ε_{emp}.
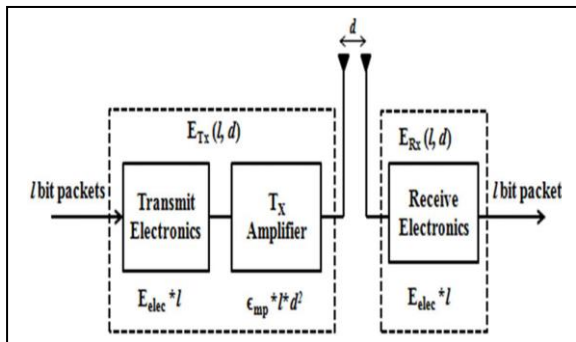

Fig. 1.   *Energy Model [4]*

$$E_{Tx} = E_{Tx\text{-}elec}(l) + E_{Tx\text{-}amp}(l, d) \tag{1}$$

$$E_{Tx} = l* E_{elec} + l * \varepsilon_{fs}\, d^2, d < d_0 \tag{2}$$

$$E_{Tx} = l * E_{elec} + l * \varepsilon_{mp}\, d^4, d > d_0 \tag{3}$$

$$E_{Rx} = l * E_{elec} \tag{4}$$

$$d_0 = (\varepsilon_{fs}/\, \varepsilon_{mp})^{1/2} \tag{5}$$

IV.   PROPTOCOL DESCRIPTION

A. Grey Wolf Optimizer

GWO is a metaheuristic technique inspired by the hierarchical leadership behavior and hunting of grey wolves suggested by Mirjalili et al. in 2014 [9]. The wolves belong to the Canidae family. Grey wolves are at the top of the food chain; They have a very strictly dominant social hierarchy shown in Fig.2. The leaders

of the group are a male and a female, called alpha. The second level called beta. The lowest ranking in the grey wolves is Omega. the wolf that is not alpha, beta or omega, called delta. In addition, group hunting is another interesting behavior of grey wolves. The main stages of hunting the grey wolves are as follows[9]:

- Exploring, pursuing and approaching prey.
- Orbiting around the prey until it stays motionless.
- Invading on the prey.

These stages shown in Fig. 3. For mathematical modeling of grey wolf hunting behavior, we use α (best solution), β (second solution) and δ (third solution) assuming they have the best knowledge of prey position. So GWO keeps the best three solutions (α, β and δ) and makes other search agents such as Omega to update their position to suit the best search agents. To hunt, a group of wolves surrounds the prey. The following equations are used to simulate the behavior of the prey [9].

$$D= |C.\, X_p(t) - X(t)| \tag{5}$$

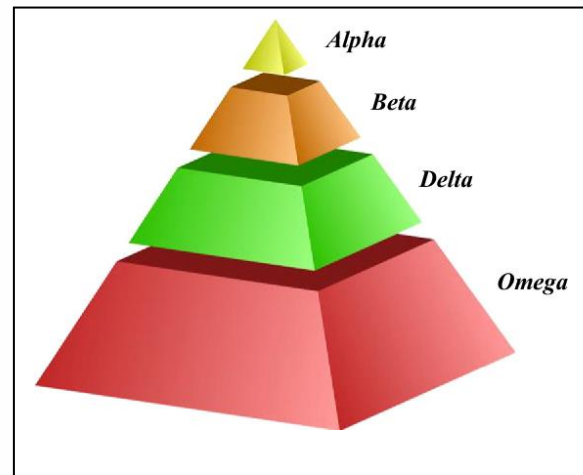$$X(t+1) = X(t) - A.\, D \tag{6}$$
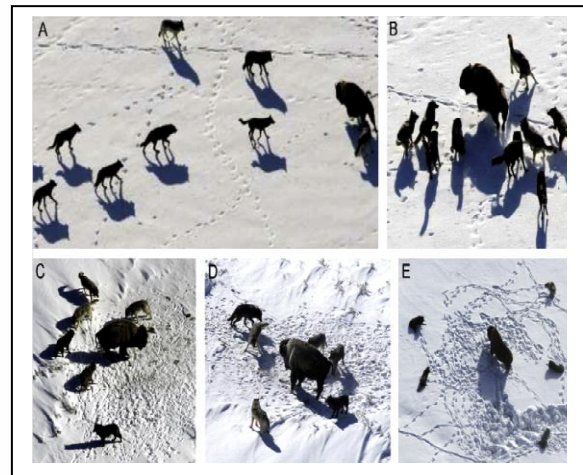

Fig. 2.   HIRARCHY OD WOLF [10]

Fig. 3.   THE MAIN STAGE OF HUNTING THE GREY WOLVES [9]

t is the current repetition, A and C are the coefficient vectors, X_p is the prey position and X is the grey wolf position. GWO uses (5) and (6) to update a search agent position. The vectors A and C are calculated by (7) and (8) [9].
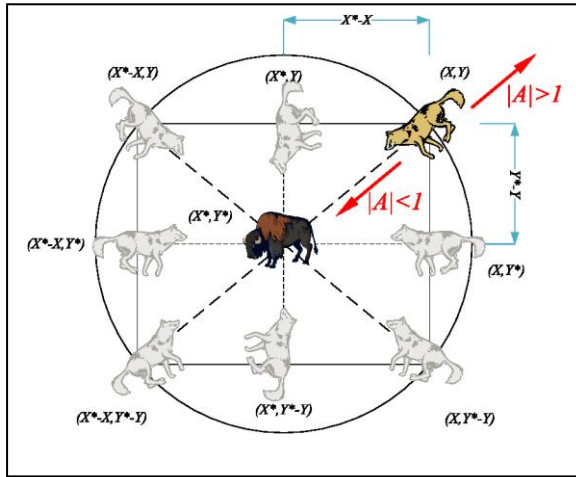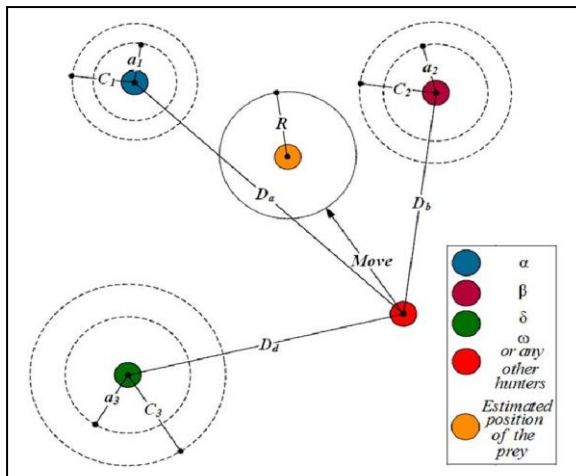

Fig. 4.   *UPDATING A SEARCH AGENT'S POSITION [9]*

Fig. 5.   *ATTACHING PREY VERSUS SEARCHING FOR PREY [37]*

$$A = 2a. r_{1} – a \qquad (7)$$

$$C = 2. r_2 \qquad (8)$$

Where the component a decreases linearly from 2 to 0 over the iteration period, r_1 and r_2 are random vectors in [0,1] [9].

$$D_\alpha= |C_1.X_\alpha – X|, D_\beta= |C_2.X_\beta – X|, D_\delta= |C_1.X_\delta – X| \quad (9)$$

$$X_1= X_\alpha –A.(D_\alpha), X_2=X_\beta –A.(D_\beta), X_3=X_\delta –A.(D_\delta) \quad (10)$$

$$X(t+1) = (X_1 + X_2+ X_3) / 3 \qquad (11)$$

Alpha, Beta, and Delta estimate the prey's position, and other wolves update their position randomly in the area around the prey using (9) and (10). Each search agent has three D and X and finally by using (11) the position of each agent updated. Fig.4. shows that the final

position is located in a random place within a circle defined by the position of α, β and δ. [9]

The grey wolves attacking the prey when it stops moving and end the hunting. To model the prey approach mathematically, we reduce the value of a. It should be noted that the range of vector A is also reduced by a. In other words, A is a random value in [-2a, 2a]. Fig. 5 [37] shows that the value of |A|<1 makes the wolves attack the prey and the value of |A|>1 makes wolves more likely to divert from prey and find a better one, besides Fig.5 shows how a grey wolf in coordinate of (X,Y) could update its next position consider to the prey position in coordinate of (X*,Y*). Different positions could obtain according to the prey position and values of A and C. Another component of the GWO that affects the exploration process is the C value. As can be seen in (8), the vector C has random values in [0,2]. This component provides random weights for hunting to intensify or weaken the effect of prey position in determining the distance in (5) and (9). This component also helps the GWO to show more random behavior during optimization and thus better exploration and avoid local trapping. It should be noted that C does not decrease linearly concerning A. We need C at all times to provide random values and to perform the identification process, not only in the initial iteration but also in the final iteration. This component is very useful in preventing to trap in local optimum, especially in the final iteration[9].

Although we have mentioned in brief the benefits of GWO in the introduction and also have explained its algorithm above, there are several advantages that make us select this algorithm over other well established optimizers:

1. The GWO is a new optimization method which overcomes the limitations such as lower tracking efficiency, steady-state oscillations, and transients as encountered in perturb and observe (P&O) and improved PSO (IPSO) techniques [39]. GWO has a position vector, so this algorithm requires less memory than an algorithm such PSO (position and velocity), besides, the algorithm keeps the best three solutions in each round of the network while the PSO keeps only one solution[10].
2. To validate the performance of the GWO, statistical measures like best, mean, worst, standard deviation, epsilon, iter and sol-iter over 50 independent runs are taken [40].
3. The GWO algorithm can reveal an efficient performance compared to other well-established optimizers [39].
4. The great advantages of GWO are that the algorithm is simple, flexible, robust and easy to implement. Also there are fewer control parameters to tune [40]. Also, the mathematical model of this algorithm is new

in its kind and allows solutions to be searched in a n-dimensional space.

5. Experimental results show the superior performance of the proposed algorithm for exploiting the optimum and it has advantages in terms of exploration [37].

B. Proposed Protocol

In this research, the nodes are divided into three categories: common nodes, cluster head nodes and relay nodes [7]. The protocol performance consists of two phases: The Clustering setup phase and the data transmission phase. Both of them are performed in each round. In the clustering setup phase, the clusters and relay nodes as a path between cluster and base station are designed and the network is built. In the data transmission phase, the cluster heads collect data from all members of the cluster and send them to the relay nodes and the data were sent from relay nodes to the base station [7]. Fig. 6 shows the network topology [7].
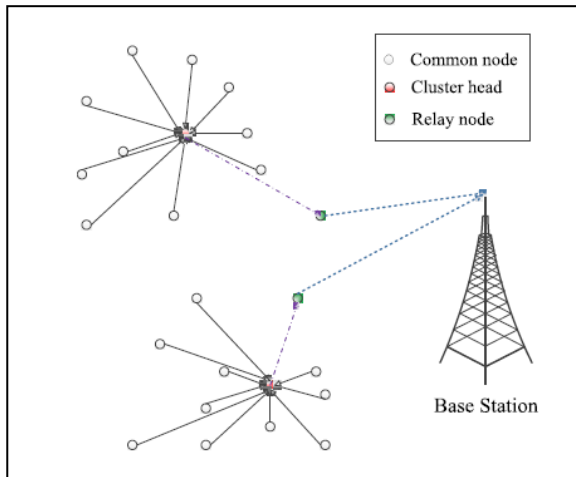

Fig. 6. *NETWORK TOPOLOGY [7]*

We assumed that N sensor nodes are divided into n clusters and randomly placed in a field, cluster heads are responsible for coordinating nodes in clusters, collecting intra-cluster information and communicating to relay nodes.

The proposed method is a centralized one, since the centralized networks have more efficiency in producing cluster quality compared with distributed networks and let clusters to be formed balanced and this caused more energy efficiency [36]. Also, to create balanced clusters, selection of the appropriate cluster heads is one of the issues that we have considered in this study. To avoid the imbalance cluster head selection, re-selecting of the cluster head seems the right solution, on the other hand, as mentioned earlier, repeating the cluster head selection increases energy consumption and ultimately reduces the network lifetime. So, for decreasing energy consumption, we will not have cluster head selection in some rounds, which causes increasing network lifetime. The Network schedule is shown in Table II. As can be seen, the clustering algorithm is repeated until the end of the network lifetime [4].

TABLE II.   NETWORK SCHEDULE [4]

Round 1	Round 2	Round 3
1st clustering	*2nd clustering*	*3rd clustering*

1) Cluster Head Selection

Cluster head selection happens in three different clustering; The details as follows:

a) First Clustering

Improving energy utilization in WSNs should be considered as a major parameter in comparison with the other algorithms (the less energy consumption, the more energy efficiency). Increasing the number of neighbors results in increasing density, therefore neighbors of each node considered as the second parameter of the cost function in first clustering. This clustering is performed in rounds of 1, 4, 7, ... So, in the first clustering "energy "and "number of neighbors" consider as cluster head selection criteria. The base station selects the cluster heads using these parameters and GWO; So it forms the clusters with uniform distribution of nodes. The cost function of the first clustering as follows:

$$\text{Cost}_1 = \alpha\, R_{energy} + (1-\alpha)\, R_{neighbors1} \quad (12)$$

As you can see, (12) consists of two parts. Constant α denotes the participation of two parts of the equation. R_{energy} is displayed in (13), that represents the ratio of the energy of all nodes $E(n_i)$ to the energy of all cluster heads $E(CH_{p,\,k})$ in the current round. n_i is represent nodes in network. $CH_{p,\,k}$ represents the cluster heads of cluster C_k of particle p.

$$R_{energy} = \Sigma\, E(n_i) / \Sigma\, E(CH_{p,k}) \quad (13)$$

$R_{neighbors1}$ is displayed in (14). Which represents the maximum average Euclidian distance of nodes to their cluster heads. $d\,(n_{ik},\,(CH_{p,\,k}))$ is the distance between n_{ik} and $CH_{p,\,k.}$ $|C_{p,\,k}|$ is the number of nodes of cluster C_k of particle p and n_{ik} represents the node i of cluster k (number of cluster). It is considered that if a node contains much residual energy as well as a larger number of neighbors it is more likely to be selected as the cluster head. The defined cost function attempts to optimize the network energy efficiency with (13) and reduces the intra-cluster communication through (14).

$$R_{neighbors1} = \max\, \{\Sigma_{\forall ni\,\varepsilon\, Cp,k}\, d\,(n_{ik},\,(CH_{p,k}))/\,|_{Cp},\kappa|\, \}(14)$$

b) Second Clustering

At the end of first clustering, since the node's energy and neighbors don't be changed, so it is more likely that the same nodes are chosen as cluster heads, therefore there is no cluster head selection in second clustering and the previous cluster heads keep their role in this clustering. This causes reducing sent and received messages.

c) Third Clustering

Third clustering is performed in rounds of 3, 6, ... Since no cluster head selection in second clustering and

energy reduction of nodes due to network activity and the importance of energy as one of the main parameters of wireless sensor network performance, the "energy of nodes" is used as one of the cost function parameters in third clustering. In addition to energy, the location of each node in the cluster is also important. So, "distance to the previous cluster head" is used as the second parameter of the cost function in third clustering. The cost function is as follows:

$$Cost_3 = \alpha R_{energy} + (1-\alpha) R_{neighbors3} \quad (15)$$

R_{energy} is similar to the first clustering which defined in (13) and $R_{neighbors3}$ is defined in (16).

$$R_{neighbors3} = \min \{\Sigma_{\forall n_i \varepsilon Cp,k} \, d \, (n_{ik}, (CH_{p,k}))/ \, |Cp,\kappa| \} (16)$$

Which represents the minimum average Euclidian distance of nodes to their cluster heads. Here, the distance between each node and its cluster head is calculated and we look for the node that has appropriate energy as well as less distance to the previous cluster head to have more number of neighbors. The selected cluster heads in third clustering have an alternative role and have been influenced by selected cluster heads in first clustering and they are not the best choices. This clustering is performed to the recovery of selected cluster heads in first clustering, that have lost so much energy.

2) Relay Node Selection

The use of relay nodes reduces the energy consumption on the cluster heads. Actually, by dedicating one relay node to each cluster head, the cluster head no longer needs extra energy for its next step [18]. A node could be chosen as a relay node if it has two criteria as follows:

1) They should have a higher energy level, due to consuming more energy than common nodes.
2) They should have a better location between the cluster head and base station.

We define the set of relay nodes as RN = {RN_1, RN_2, …, RN_m} and the set of common nodes as CN [7]. Similar to the cluster head selection section, we use the cost function to select relay nodes with the difference that we use one cost function for all clustering. The cost function is as follows:

$$Cost = \beta R_{energy} + (1-\beta) R_{distance} \quad (17)$$

R_{energy} represents the ratio of the average energy of relay nodes E_{RN} to average energy of common nodes E_{CN}, which is defined in (18).

$$R_{energy} = E_{RN} / E_{CN} \quad (18)$$

E_{RN} (z) defines the energy of relay nodes and E_{CN} (z) defines the energy of common nodes. |RN| and |CN| are the number of relay nodes and common nodes, respectively. $R_{distance}$ represents the minimum average Euclidian distance of relay nodes to the BS. This

equation means that the minimum distance to the BS is suitable for a relay node. Which is defined as follows:

$$R_{distance} = \min \{\Sigma_{i=1} \, d \, (RN_i, (BS))/ \, |PN| \} (19)$$

C. Clustering Formation

1) Clustering Set Up Phase

In wireless sensor networks, each node is assigned an index (ID) according to its location. The selection of relay nodes and cluster heads is performed by BS. This phase is as follows [7]:

- At first, each node sends a NODE-MSG message to broadcast the information of its energy and location, which is necessary for the selection of relay nodes and cluster heads.

- Then the BS selects the cluster head with (12). In second clustering have no cluster head selection and in third clustering, the BS selects the cluster head with (15). Then BS broadcasts a message consists of a cluster head index to inform the network from cluster head location. after each cluster head knows its conditions, it introduced itself to the network by a CH-ADV message. This message contains a cluster head index and a header which identified it as an advertisement message.

- Then, similar to the cluster head selection, the BS selects the relay node using (17). When a relay node is chosen, the BS sends an advertisement message (for example, RN-ADV) that contains node index, a cluster head index and a header to the network to inform their conditions as a relay node. Each common node chooses its cluster that needs less energy to transmit information.

- After each common node decides which cluster wants to be join, it informs the cluster head by sending the JOIN-REQ message. The message is too short, which contains a node index, a cluster head index and the residual energy of the node. In this way, clusters formed and the duty of each node determined.

The cluster head acts as a control center to transmit data. The cluster head sets up the TDMA schedule and broadcasting a SCHEDULE-MSG message to all nodes in the cluster and relay node. This avoids the data collision and also allows the radio component of each common node and relay node to be switched off all the time except in sending and receiving time. This causes saving more energy. When the TDMA schedule is known to all common nodes the clustering setup phase completed and the data transmission phase begins at the same time.

2) Data Transmission Phase

In this phase, the common nodes send their data to their cluster head by TDMA schedule. All nodes synchronized with the synchronization pulse of the BS. Cluster head should always awake to receive data from each common node. Then, the cluster head sends aggregated data to the relay node. By TDMA

scheduling that is run by cluster head, sensor nodes can switch on/off the radio component to save energy.

V. SIMULATION AND ANALYSIS

The proposed method is analyzed through simulation in MATLAB software. The proposed algorithm has been compared with LEACH [21], LAN [14], EACCC [13], PSO [22], AOSTEB [19] algorithms under the same operating conditions. "Network lifetime", "Number of dead nodes in each round", "Total remaining energy (TRE)" in cluster head nodes and relay nodes are the evaluation criteria. Simulations have been studied in three different scenarios with various conditions. We considered the effect of number of nodes, network dimensions, and BS location to investigate the scalability of our work in our simulations.

A. Scenario 1

The simulation parameters of this scenario are illustrated in Table III.

TABLE III.   SIMULATION PARAMETERS OF SCENARIO1

Parameter	Value
Simulator	MATLAB
Electronics energy	50 nJ/bit
Initial energy	0.5J
Energy for data aggregation (E_{DA})	5 nJ/bit/signal
Communication energy (εfs)	10 pJ/bit/m2
Communication energy (εamp)	0.0013 pJ/bit/m4
Packet size	6400
Sensing area (M × M)	100 m ×100 m
Number of sensor nodes (N)	100
Base station location	(50 ,50)

Fig. 7, illustrates the results of comparing the algorithms in terms of "Network lifetime" that have three parameters called: FND (first node died), HND (half node died), and LND (last node died). In the proposed method, after selecting CHs in first clustering, the same nodes (if alive) will be selected as the cluster heads of second clustering. Even because these nodes are suitable, they may be selected as the cluster heads of the third clustering, this, leads them to death because the energy is reduced more in these nodes than the others. This is why comparing the FND of the proposed method with the other methods has not ideal results. On the other hand, the selection of other nodes as cluster heads in subsequent clustering avoids the drastic reduction of energy, which results in the increasing energy in the middle and final rounds. HND and LND parameters confirms these results. As can be seen, the proposed method has a better performance than the classical and metaheuristics algorithms. Reducing the number of clustering and decreasing the energy consumption of cluster heads due to the use of relay nodes, considering "energy" and "number of neighbors" as cost function criteria and using the GWO to find the most optimal CHs, provides an appropriate approach that results in increasing network lifetime." The number of dead nodes in each round" which is shown in Fig. 8, is the second criterion that considered

in this work. The result illustrates that the proposed method has less dead nodes than the other methods in each round, the most important reason could be considered the type of cluster head selection.
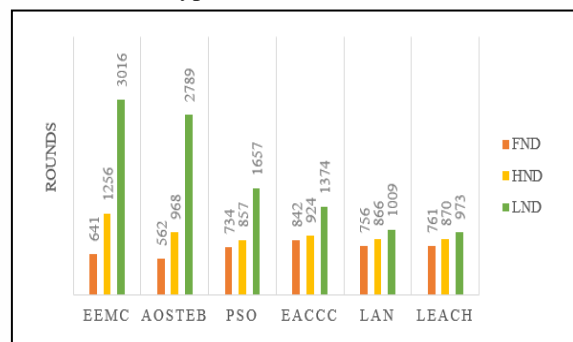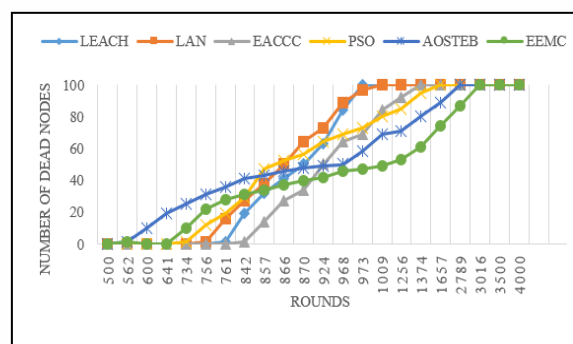

Fig. 7.   Network Lifetime in scenario1

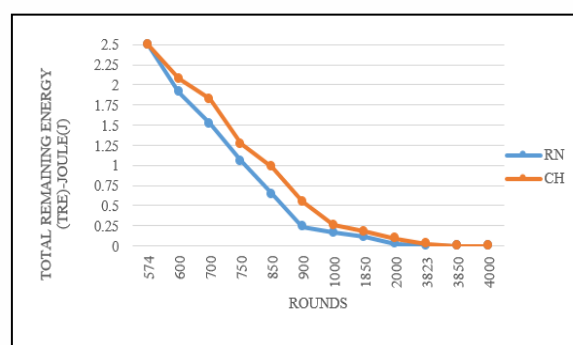Fig. 8.   Number of dead nodes in each round

Fig. 9.   Total remaining energy (TRE) of the cluster head and relay nodes

The "Total remaining energy (TRE) of cluster head and relay nodes" is the third criterion to be examined. The results are shown in Fig. 9. As mentioned, having more energy than other nodes, was one of the criteria for selecting these nodes in the network, and also the main reason for the presence of relay nodes in the network was to deplete the energy consumption in the cluster heads. As you can see in Fig. 9, the utilized energy in the relay nodes has been reduced more than in the cluster head nodes and this has prevented the energy depletion in the cluster heads.

B. Scenario 2

In this scenario, the comparisons are based on the parameters in Table IV. By doubling the network dimension and number of nodes and changing the BS location, we also investigate the scalability of the

network. As in the previous scenario, "Network lifetime", "Number of dead nodes in each round", "Total remaining energy (TRE) in the cluster head and relay nodes" have been investigated. The first parameter considered in this scenario is "network lifetime".

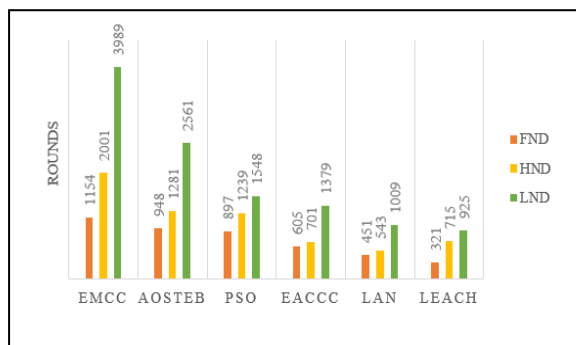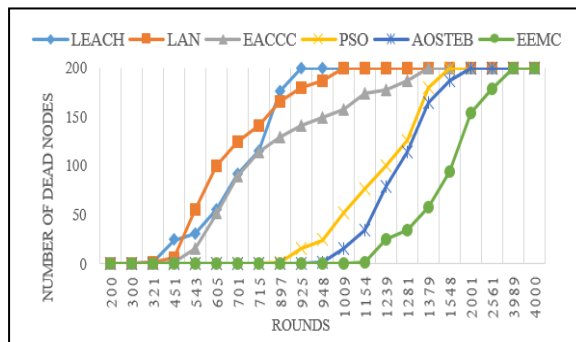| Parameter	Value
Simulator	MATLAB
Electronics energy	50 nJ/bit
Initial energy	1J
Energy for data aggregation (E_{DA})	5 nJ/bit/signal
Communication energy (εfs)	10 pJ/bit/m2
Communication energy (εamp)	0.0013 pJ/bit/m4
Packet size	6400
Sensing area (M × M)	200 m ×200 m
Number of sensor nodes (N)	200
Base station location	(100 ,200)

Fig. 10. *Lifetime in scenario2*

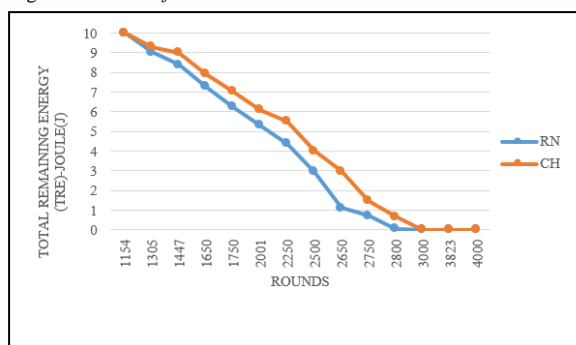Fig. 11. *Number of dead nodes in each round*

Fig. 12. *Total remaining energy (TRE) of the cluster head and relay nodes*

According to Fig. 10, the proposed method improved HND and LND by selecting the best CHs, as mentioned earlier, since the nodes selected as the cluster head may perform this role for two or three successive rounds, the FND values may not be ideal but it's still in better conditions than the other algorithms. The "Number of dead nodes in each round" is the second parameter considered in this scenario. Fig. 11, shows the results of the comparison of different algorithms. As could be seen the proposed method enjoys the less dead nodes in each round than the other algorithms. This due to the decreasing selection of cluster heads in some rounds, which caused the reduction of sent and received messages. "Total remaining energy (TRE) of the cluster head and relay nodes" is the last parameter examined, its results shown in Fig. 12. We consider the initial energy in this scenario as 1 joule, as we all know, in this paper, the cluster head acts as a control center and has a lot of duties and all these tasks require energy. So we have used relay nodes due to eliminating utilized energy of cluster heads in finding their next step. It is obvious that this, reduces the energy consumption in the cluster heads. The results shown in Fig. 12, confirms these findings. By examining the results in this scenario we found that by increasing the number of nodes and network dimensions twice as the first scenario, the FND, HND, and LND have improved by 1.8, 1.6 and 1.3 times, respectively. These results are due to the use of GWO, which is a flexible and scalable algorithm.

C. Scenario 3

In this scenario, the comparison parameters are shown in Table V.

Parameter	Value
Simulator	MATLAB
Electronics energy	50 nJ/bit
Initial energy	5J
Energy for data aggregation (E_{DA})	5 nJ/bit/signal
Communication energy (εfs)	10 pJ/bit/m2
Communication energy (εamp)	0.0013 pJ/bit/m4
Packet size	6400
Sensing area (M × M)	500 m ×500 m
Number of sensor nodes (N)	500
Base station location	(500 ,500)

By increasing the number of nodes and dimensions of the working environment in about 5 times we compared the scalability of the proposed method. Similar to the two previous scenarios, "Network lifetime", "Number of dead nodes in each round" and "Total remaining energy of cluster head and relay nodes" are investigated. The first compared parameter is "Network lifetime" shown in Fig. 13. The proposed method improved LND, and HND. Through increasing the number of nodes and dimensions of the working environment, it is observed that the proposed method shows better performance and this is due to the flexibility of the GWO. By examining the results, we found out the "Network lifetime" comparison with the first and second scenarios, have improved by 2.1 and 1.5 times, respectively. What stands out from Fig. 13 is that the proposed method has better performance

than the classical and metaheuristic algorithms. This is caused by type of CH selection and decreased consuming energy due to the use of R. The "Number of dead nodes in each round" is the second criterion. The result is shown in Fig. 14. The findings illustrate that the proposed method has fewer dead nodes than the other methods, the most important reason could be the type of cluster head selection. We reduced the energy consumption in each round by reduction in number of clustering and this caused the lessen dead nodes in each rounds. Fig. 15, shown the results of the last compered parameter, "Total remaining energy (TRE) of cluster head and relay nodes". Since the cluster heads should be always on in the network, of course they lost much more energy than other nodes so we used the relay nodes to decline the energy dissipation in cluster heads and Fig. 15, illustrates that our goal obtained.
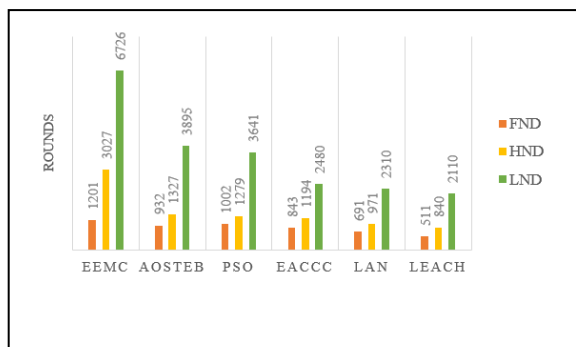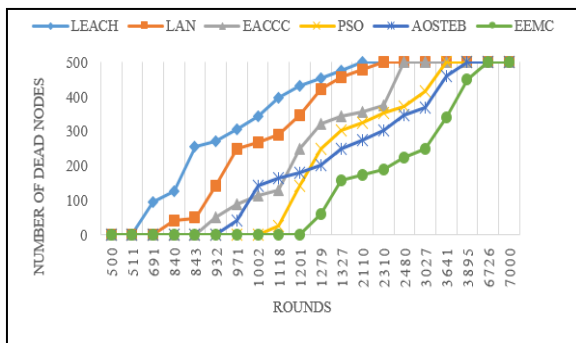

Fig. 13. *Network Lifetime in scenario3*

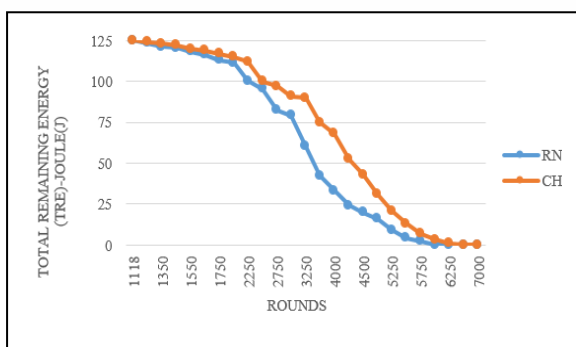

Fig. 14. *Number of dead Nodes in each round*

Fig. 15. *Total remaining energy (TRE) of the cluster head and relay nodes*

VI. CONCLUSION

In this paper we provide a new centralized clustering using GWO, and have improved network lifetime by not having cluster head selection in some rounds that caused by reducing the number of sent and received messages. Besides, we used relay nodes so the cluster heads would not consume energy to find their next hop. The proposed method has defined some cost functions to select the cluster head and relay nodes using GWO. We compared proposed method in three scenarios with other algorithms in terms of "Network lifetime" and "Number of dead nodes in each round"; also we compared the "Total remaining energy (TRE) of cluster head and relay nodes" in each scenario. The proposed method contains three different clustering with various conditions. Since the nodes distribution and distance to BS affect in the cluster heads selection, so we considered the impact of different conditions such as "network dimension", "number of nodes" and "BS location" in our investigations. In the first of them, the BS is placed at the center of working place, in the second and third one, the BS is placed in the margins of the working place. By investigating the "network lifetime" we found out that the proposed method improved the values of LND and HND in all scenarios, but the FND has not the ideal values (this will be investigated in our future work), also the proposed method has the fewer dead nodes than others. The performance of the last scenario in comparison with first and second scenarios has improved by 88% and 60%, respectively, and It is due to the scalability of the GWO. "Grey Wolf Optimizer", "relay nodes", "no cluster head selection in some rounds" and "a centralized network" have considered as the innovations of our research.

We are suggesting this article for the centralized networks which want to improve the values of HND, LND, and fewer dead nodes; However, the distributed nature of wireless sensor networks made us to investigate a distributed version of this research in our future work.

REFERENCES

[1] J. N. Al-Karaki and a E. Kamal, "Routing techniques in wireless sensor networks: a survey," IEEE Wirel. Commun., vol. 11, no. 6, pp. 6–28, 2004.

[2] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," Ad Hoc Networks, vol. 3, no. 3, pp. 325–349, 2005.

[3] V. Kumar, J. K. Chhabra, and D. Kumar, "Grey Wolf Algorithm-Based Clustering Technique," J. Intell. Syst., vol. 26, no. 1, 2016.

[4] M. Mirzaie and S. M. Mazinani, "MCFL: an energy efficient multi-clustering algorithm using fuzzy logic in wireless sensor network," Wirel. Networks, pp. 1–16, 2017.

[5] S. Mirjalili, J. S. Dong, and A. Lewis, Nature- Inspired Optimizers. Springer, Cham, 2019.

[6] S. K. Shandilya, S. Shandilya, and A. K. Nagar, Advances in Nature-Inspired Computing and Applications, 1st ed. Springer International Publishing, 2019.

[7] Y. Zhou, N. Wang, and W. Xiang, "Clustering Hierarchy Protocol in Wireless Sensor Networks Using an Improved PSO Algorithm," IEEE Access, vol. 5, no. c, pp. 2241–2253, 2017.

[8] M. Ramadas and A. Abraham, "Metaheuristics and Data Clustering," in Metaheuristics for Data Clustering and Image Segmentation, 2019, pp. 7–55.

[9] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," Adv. Eng. Softw., vol. 69, pp. 46–61, 2014.

[10] H. Faris, I. Aljarah, M. Azmi, and A. S. Mirjalili, "Grey wolf optimizer : a review of recent variants and applications," Neural Comput. Appl., vol. 30, no. 2, pp. 413–435, 2017.

[11] S. J. Nanda and G. Panda, "A survey on nature inspired metaheuristic algorithms for partitional clustering," Swarm and Evolutionary Computation, vol. 16. Elsevier, pp. 1–18, 2014.

[12] M. Panda and B. Das, "Grey Wolf Optimizer and Its Applications : A Survey," in Proceedings ofthe Third International Conference on Microelectronics, Computing and Communication Systems, Lecture, 2019.

[13] P. Singh and M. Satvir, "Artificial bee colony metaheuristic for energy-efficient clustering and routing in wireless sensor networks," Soft Comput., vol. 21, no. 22, pp. 6699–6712, 2016.

[14] A. R. Pagar and D. C. Mehetre, "A Survey on Energy Efficient Sleep Scheduling in Wireless Sensor Network," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 5, no. 1. pp. 557–562, 2015.

[15] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in Proceedings of the 33rd Hawaii International Conference on System Sciences, 2000, vol. 1, p. 10.

[16] R. Sharma, V. Vashisht, A. V. Singh, and S. Kumar, "Analysis of Existing Clustering Algorithms for Wireless Sensor Networks," in System Performance and Management Analytics, P. K. Kapur, Y. Klochkov, A. K. Verma, and G. Singh, Eds. Singapore: Springer Singapore, 2019, pp. 259–277.

[17] A. Manjeshwar and D. P. Agrawal, "TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Network," in Proceedings 15th International Parallel and Distributed Processing Symposium. IPDPS 2001, 2001.

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

[19] S. Lindsey and C. S. Raghavendra, "PEGASIS: Power-efficient gathering in sensor information systems," in Proceedings, IEEE Aerospace Conference, 2002, vol. 3, pp. 1125–1130.

[20] O. Younis and S. Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks," IEEE Trans. Mob. Comput., vol. 3, no. 4, pp. 366–379, 2004.

[21] N. Al-humidi and G. V Chowdhary, "Energy-Aware Approach for Routing Protocol by Using Centralized Control Clustering Algorithm in Wireless Sensor Networks," in Computing, Communication and Signal Processing, 2019, pp. 261–274.

[22] Y. Zhong, Z. Huang, and Y. Xu, "An Energy Optimal Clustering Algorithm," J. Phys. Conf. Ser., vol. 1229, 2019.

[23] M. S. Azizi and M. L. HASNAOUI, "An energy efficient clustering protocol for homogeneous and heterogeneous wireless sensor network," 2019.

[24] T. Ahmad, M. Haque, and A. M. Khan, An Energy-Efficient Cluster Head Selection Using Artificial Bees Colony Optimization for Wireless Sensor Networks. Springer International Publishing, 2019.

[25] S. Z. Selim and K. Alsultan, "A SIMULATED ANNEALING ALGORITHM FOR THE CLUSTERING PROBLEM," Pattern Recognit., vol. 24, no. 10, pp. 1003–1008, 1991.

[26] M. Dorigo and G. Di Caro, "Ant colony optimization: A new meta-heuristic," in Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406), 1999, vol. 2, pp. 1470–1477.

[27] R. Eberhart and J. Kennedy, "A New Optimizer Using Particle Swarm Theory," in MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 1995, pp. 39–43.

[28] J. H. Holland, "Genetic Algorithms," Sci. Am., vol. 267, no. 1, pp. 66–73, 1992.

[29] S. Sharma and E. A. Bindle, "Comparison of Routing Algorithm of Optimized and Non-Optimized PSO in Wireless Sensor Networks," Int. J. Eng. Sci. Comput., vol. 9, no. 3, pp. 1–6, 2019.

[30] Y. Shi and R. Eberhart, "A Modified Particle Swarm Optimizer," in 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360), 1998, pp. 69–73.

[31] P. S. Shelokar, V. K. Jayaraman, and B. D. Kulkarni, "An ant colony approach for clustering," Anal. Chim. Acta, vol. 509, no. 2, pp. 187–195, 2004.

[32] D. KARABOGA, "An idea based on honey bee swarm for numerical optimization," Tech. report-tr06, Erciyes Univ. Eng. Fac. Comput. Eng. Dep., vol. 200, pp. 1–10, 2005.

[33] P. Nayak, "Genetic Algorithm Based Clustering Approach for Wireless Sensor Network to Optimize Routing Techniques," in 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, 2017, pp. 373–380.

[34] D. Reddy, E. Mahesh, C. Kongara, and R. Cheruku, "A PSO Based Routing with Novel Fitness Function for Improving Lifetime of WSNs," Wirel. Pers. Commun., no. 1/2019, 2018.

[35] V. K. Arora, V. Sharma, and M. Sachdeva, "ACO optimized self - organized tree - based energy balance algorithm for wireless sensor network," J. Ambient Intell. Humaniz. Comput., vol. 0, no. 0, p. 0, 2019.

[36] P. S. Mann and S. Singh, "Energy efficient clustering protocol based on improved metaheuristic in wireless sensor networks," Netw. Comput. Appl., vol. 83, no. c, pp. 40–52, 2017.

[37] S. Mirjalili, "How effective is the Grey Wolf optimizer in training multi-layer perceptrons," Appl. Intell., vol. 43, no. 1, pp. 150–161, Jul. 2015.

[38] A. Chandanse, P. Bharane, S. Anchan, and H. Patil, "Performance Analysis of LEACH Protocol in Wireless Sensor Network," in 2nd International Conference on Advances in Science & Technology (ICAST-2019), 2019

[39] Garg, A. and Kumar Agarwal, S. (2012). Dynamic Stability Enhancement of Power Transmission System Using Artificial Neural Network Controlled Static Var Compensator. International Journal of Computer Applications, 53(9), pp.21-29.

[40] Energy-Efficient Routing Protocol for Wireless Sensor Networks Based on Improved Grey Wolf Optimizer. (2018). KSII Transactions on Internet and Information Systems, 12(6).

Maryam Ghorbanvirdi was born in Birjand, Iran, on 25 January 1991. She received her Bachelor degree in Electrical Engineering in Telecommunication orientation from Birjand University, Birjand, Iran in 2014 and her Master degree in Telecommunications, System orientation from Imam Reza International University, Mashhad, Iran in 2019.Her research interests include Wireless Sensor Networks and Internet of Things.

Sayyed Majid Mazinani was born in Mashhad, Iran on 28 January 1971. He received his Bachelor degree in Electronics from Ferdowsi University, Mashhad, Iran in 1994 and his Master degree in Remote Sensing and Image Processing from Tarbiat Modarres University, Tehran, Iran in 1997. He worked in IRIB from 1999 to 2004. He also received his Ph.D. in Wireless Sensor Networks from Ferdowsi University, Mashhad, Iran in 2009. He is currently assistant professor at the faculty of Engineering in Imam Reza International University, Mashhad, Iran. He was the head of Department of Electrical and Computer Engineering from 2009 to 2012. His research interests include Computer Networks, Wireless Sensor Networks and Smart Grids.