Energy-Efficient Clustering Algorithm in Underwater Sensor Networks Based on Fuzzy C Means and Moth-Flame Optimization Method

WANG FEI1, BAI HEXIANG1, LI DEYU1, AND WANG JIANJUN2
1School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China
2Grain and Oils Information Center of Jiangsu Province, Nanjing 210008, China
Corresponding author: Wang Fei (wfei@sxu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 41871286, in part by the National Natural Science Foundation of China 61672331, and in part by the Key R&D program of Shanxi Province (International Cooperation, 201903D421041).

ABSTRACT Underwater sensor networks (UWSN) often suffers from the irreplaceable batteries and high delay of long-distance communications, thus one of the most important issues on UWSN is how to extend the lifespan of the network and balance the energy consumption of each node by reducing the transmission distances. Actually, clustering method is one of the main methods to resolve the problem. In the clustered UWSN, the major concerns are obtaining appropriate number of clusters, forming the clusters and selecting an optimal cluster head(CH) with each cluster. This paper proposes a novel hybrid clustering method based on fuzzy c means (FCM) and moth-flame optimization method (MFO) to improve the performance of the network (FCM/MFO). The idea is to form energy-efficient clusters by using FCM and then use an optimization algorithm MFO to select the optimal CH within each cluster. The simulation results validate the energy-efficient performance of FCM/MFO in comparison with the other existing algorithms. The results clearly show the significant impact of FCM/MFO on energy-efficiency in UWSN.

INDEX TERMS UWSN, clustering algorithm, fuzzy C means, moth-flame optimization.

I. INTRODUCTION
Being self-organized and easily-deployed, underwater sensor network (UWSN) is suitable for target detection, target tracking and parameter monitoring in large-area waters [1]. However, UWSN faces two challenges 1) Sensor nodes have limited energy and will consume more energy when the underwater acoustic transducer used by them sends data packets. Moreover, the energy consumption increases with transmission distance. 2) Compared with the terrestrial sensor networks, UWSN has defects in large propagation delay, low bandwidth, high bit error rates and so on. Therefore, UWSN is not suitable for long-distance transmission [2], [3].

According to the above, network life cycle, energy consumption and communication distance are 3 main considerations for designing UWSN. Topology control using the clustering method is one of the main solutions to the problems of the UWSN, because it can decrease the communication interference, balance the energy consumption and prolong the lifetime of the network. The purpose of clustering is to divide the entire underwater sensor networks into multiple regions. In each region, sensor nodes only communicate with the cluster head within their own cluster. This new topology structure reduces the energy consumption, the interference and attenuation among the sensor nodes, when they send the data packets for long distances. At the same time, it prolongs the life cycle and improves the networks’ performance [4], [5]. Soft computing technologies such as FCM and MFO are considered as the important methods to solve the clustering problem by providing better topology control which could reduce the energy consumption of the networks.

In FCM method, the samples are classified according to their memberships, and this method is more likely to reflect the real-world. The node clustering process of UWSN is similar to the data clustering process of FCM, so the clustering of UWSN can be mapped to fuzzy clustering partition of sample space. In FCM, each node can be considered as a sample of the network, simultaneously, the nodes close to the center of the cluster are the cluster heads, and the subsets obtained by the clustering are clusters in UWSN.

MFO algorithm treats moths as candidate solutions where the locations of moths in space are considered as the variables.
for the problem. Through a fitness function, moths search optimal solutions in space, which are considered as flames in MFO. Thus, the optimal locations of cluster heads can be considered as flames. Therefore, we can map the election of the cluster heads to the solving of MFO algorithm. In this paper, we propose a hybrid method for clustering and selecting of cluster heads based on FCM and MFO. This method can effectively reduce the energy consumption of the networks and balance the load of the sensor nodes.

II. RELATED WORK
A. PROGRESS OF CLUSTERING TECHNOLOGIES

Saving energy is an important objective in the application of UWSN, and clustering method is a well-known technology used for lowering energy consumption of UWSN. There is a lot of research work on clustering and selection of cluster heads for minimizing the UWSN energy consumption. In this section, some of the research work is discussed.

In literature [6], the authors proposed Low-Energy Adaptive Clustering Hierarchy (LEACH). It is a classical algorithm in the hierarchical networks, which selects the cluster heads randomly in a cyclic manner with an even distribution of energy to the load. The LEACH algorithm enables efficient energy consumption, but disregards the information of the nodes, when selecting the CH nodes, i.e., residual energy, the energy consumption of communication and the number of the neighbor nodes. In an improved version of the LEACH algorithm, LEACH-C, all of the sensor nodes transmit the information, such as location and residual energy to the base station (BS). Moreover, LEACH-C, used as centralized algorithm, reduces the energy consumption of each node in the network [6]. However: LEACH-C suffers from some defects. For example, in LEACH-C, each node has equal possibility to be a CH. As the energy of the network becomes low, the nodes whose residual energy is low should also be selected to be CHs, which will result in energy inefficiency and unbalanced energy load.

Particle swarm optimization (PSO) based techniques are another popular technique for routing algorithms in UWSN. Singh et al. [7] proposed energy aware CH selection algorithm using PSO, which constructs the fitness function by introducing the residual energy, distance and node density. However, this algorithm does not consider the cluster formation process, resulting in energy inefficiency of the entire network.

Rao et al. [8] proposed a PSO-based energy efficient cluster head selection (PSOECHS) to solve the problem in [7], which takes the distances to the base station (BS) into consideration and adds this condition to the fitness function. PSOECHS extends the lifetime of the network by varying the way for the CHs selection and allows the adjustment of the number of the member nodes of each cluster. However, the energy consumed by each node for computing and selecting CHs is high.

Zhang et al. [9] proposed clustering architecture for medium scale UWSN using discrete PSO and genetic algorithm (GA) to prolong the network lifetime. However, the stability of the algorithm is poor and the clustering model is not completely fit for UWSN. To resolve the drawbacks in [8] and prolong the network lifetime further, Pengwei Li et al. [10] proposed an improved particle swarm optimization Algorithm of clustering in UWSN, which take into account the required transmission power of nodes. However, the particle coding is complex for the sensor nodes. The literature [11] proposed Energy Centers Searching using PSO(EC-PSO) to avoid these energy holes and search energy centers for CHs selection. In addition, random reinitialization was used to avoid CHs getting too close and a protection mechanism with threshold value was utilized to keep the low energy nodes from forwarding. However, this algorithm completely ignores the distance among nodes, CHs and sink node when constructing fitness function, thus impacts the energy consumption.

Fuzzy enable clustering technique was also used for optimizing the routing algorithm. In literature [12], a decentralized fuzzy c-means energy-efficient routing protocol has been proposed to minimize total network energy and effectively prolong the network lifetime. However, the election of CHs is done locally in each cluster and this algorithm does not utilize optimization technique. In literature [13], Bhatti proposed fuzzy c-means clustering and energy-efficient CH selection for cooperative sensor networks to save energy, in which clusters are formed using FCM method and CHs is selected based on the parameters like location, SNR and residual energy of the nodes. However, the topological structure of the sensor network is two-dimensional space. The literature [14] proposed a balanced energy consumption clustering routing protocol for a wireless sensor network (BECR). In BECR, the improved FCM method was first used to divide the network into clusters, and the node closest to the cluster center was regarded as the initial cluster head. When the energy of the initial cluster head dropped by 20%, the fuzzy logic system (FLS) was then used to select the cluster head for each cluster. When the energy of the cluster head each dropped by 10%, the next round of cluster head elections was performed, otherwise, the cluster head election was not performed. However, this algorithm suffers from the random problem, and it’s easy to fall into local optimality.

Till date, various researchers have used the hybrid method, which helps in improving the shortcoming of one approach and fusing the advantages of these approaches. The literature [15] used FCM for the cluster creation and then selected the best node as a CH within each cluster formed. By using an evolutionary technique DE, for the CH selection, the fitness of each node was calculated through a designed fuzzy inference system. The literature [16] utilized artificial bee colony algorithm to adjust the fuzzy rules of LEACH-SF. The fitness function of the algorithm was defined to prolong the network lifetime, based on the application specifications. This algorithm used LEACH protocol and fuzzy C-means method to form balanced clusters and prolong the lifetime of the network. However, the proposed clustering technique
was designed for the networks with stationary sensor nodes. The literature [17] used fuzzy clustering to join sensors in the network and the PSO algorithm to estimate the initial value of cluster heads. First, it used fuzzy technique for clustering, then selected the optimal cluster head by applying PSO to fuzzy table. The algorithm effectively reduced the energy consumption. In addition, there exists some clustering methods like differential evolution and simulated annealing [18], fuzzy c-means and genetic fuzzy system [19]. Table 1 represents categorizations of clustering method along with clustering objectives and disadvantages.

As the works mentioned above, there have been many studies on clustering technologies for UWSN. In this paper, we expect to provide important support for designing a hybrid clustering method based on FCM and MFO to improve the performance of the UWSN.

In MFO, the moths move around, and flames are considered as pins that are dropped by moths when searching the search space. Therefore, each moth searches around a flame and updates it in case of finding a better flame location. Therefore, the moth will always find the best solution by using MFO. To the best of our knowledge, it is the first paper to introduce the clustering algorithm of UWSN based on a hybrid method which implements FCM and MFO.

B. OUR CONTRIBUTIONS

1) To the best of our knowledge, this is the first work that applies the hybrid method with FCM and MFO algorithms to improve the performance of UWSN.

2) We use FCM to cluster the network, and use MFO to elect the optimal cluster heads. In the proposed method, we use elbow method to determine the optimal number of the clusters, and implement FCM to decide the locations of the initial cluster heads.

3) We construct the fitness function after taking into consideration of the following: the distances between nodes and candidate cluster heads, candidate cluster heads, sink node, and energy consumption of the network. Then MFO is utilized to elect cluster heads through the fitness function.

4) To evaluate the performance of our proposed method, we compared it with some similar methods, such as LEACH, FCMPSO and EEFCM-DE.

III. CLUSTER INITIALIZATION

In this section, we divide the sensor nodes into several clusters initially. During the formation of the clusters, the nodes closer to their own cluster centers and requiring less energy for transmitting data packets should be assigned to the same cluster.

In this paper, we propose a method to divide the networks into c clusters initially, and each of which has an initial clustering center. Specifically, each node should find their clusters by the parameters called membership degree uij. That is to say, if node i satisfies the below condition, it belongs to the cluster j.

\[ i = \text{argmax}(u_{ij}) \quad (1) \]

In addition, we can find the optimal number c of the clusters by elbow method.

A. PROPOSED CLUSTERS INITIALIZATION METHOD

Most clustering methods initialize the cluster center randomly. However, these methods have a certain probability that the clustering center election falls into a local optimum, resulting in cluster head non-uniform. This paper initializes the clustering center by taking the distances between nodes and the energy consumption of nodes into consideration.

Firstly, we can use equation (2) to determine the number k of cluster heads.

\[ k = \sqrt{n/2} \quad (2) \]

n is the amount of the sensor nodes in the networks, from which k nodes are selected as initialization cluster centers.

\[ Y = (E_r/E)^{-1} \|X_i - X_j\| \quad (3) \]

E is the total energy of the sensor nodes, and \|X_i - X_j\| is the distance between nodes i and j. The smaller Y is, the more possibly the node will be selected as the initialization cluster center.

Secondly, we will calculate the membership degree of each node, then we use elbow method to determine the optimal number of the clusters.

Finally, we can obtain the locations of the initial cluster heads by iteration.

1) CONSTRUCTION OF MEMBERSHIP MATRICES

Membership degree is a function that represents the degree to which the object x belongs to the set A. The variable of the function is all the objects in the set A, and the value range is \([0,1]\). In this paper, membership degree mainly refers to the distances from node to the clustering center. The calculation formula of membership degree is as follows [20]:

\[ u_{ij} = \left[ \sum_{k=1}^{C} \left( \frac{\|X_i - V_j\|^2}{\|X_i - V_k\|^2} \right)^{1/2} \right]^{-1} \quad (4) \]

uij is the members of the membership matrix. When the membership matrix is obtained, the clustering centers can be updated, and the update formula is as follows:

\[ V_j = \frac{\sum_{i=1}^{N} u_{ij}^m X_i}{\sum_{i=1}^{N} u_{ij}^m} \quad (5) \]

The values of the clustering center before and after updating are compared as follows:

\[ \|V_a - V_b\| < \varepsilon \quad (6) \]

If formula (6) is true, the iteration will be stopped. Otherwise, if it is not true, formula (4) and (5) will be reran to update the cluster center.
| Algorithm            | Method Type | Clustering Objective                                    | Disadvantages                                                                 |
|----------------------|-------------|----------------------------------------------------------|-------------------------------------------------------------------------------|
| LEACH[6]             | probabilistic | minimizing energy consumption | does not consider information of nodes when selecting CHs to cause network energy imbalance |
| LEACH-C[6]           | probabilistic | improving network lifetime | results in energy inefficiency and unbalanced energy load for equal possibility to elect CHs |
| EACHS-PSO[7]         | heuristic    | energy-efficient CHs election to improve network lifetime | does not consider the cluster formation process |
| PSOECHS[8]           | heuristic    | create efficient clusters and select CHs for energy efficiency | increases energy dissipation when selecting CHs |
| RCA-UASN[9]          | heuristic    | create energy-efficient clustering | the stability of the algorithm is poor and the architecture in medium scale UWSN clustering model is not completely fit for UWSN |
| IPSOC-UASN[10]       | heuristic    | to lengthen the network lifespan by improved PSO | the particle coding is complex for the sensor nodes |
| EC-PSO[11]           | heuristic    | to avoid energy holes and search energy-efficient CHs | ignores the distance among nodes, CHs and sink node when constructing fitness function |
| DFCMER[12]           | non-deterministic | to prolong the network lifetime by a FCM based protocol | the election of CHs is done locally in each cluster and this algorithm does not utilize optimization technique |
| FCM-EECHS[13]        | non-deterministic | to form energy-efficient clusters to save energy | the topological structure of the sensor network is two-dimensional space |
| BECR[14]             | non-deterministic | to balance energy consumption and prolong the lifetime | suffers from the random problem, and it’s easy to fall into local optimality |
| EECFM-DE[15]         | Hybrid       | Energy-efficient clustering to improve network lifespan and throughput | nodes consume more energy for this algorithm is distributed for each sensor |
TABLE 1. (Continued.) Categorizations of clustering algorithms, clustering objectives and disadvantages.

| Algorithm        | Type   | Objective                                      | Method                                                                 |
|------------------|--------|------------------------------------------------|------------------------------------------------------------------------|
| OFC[16]          | Hybrid | to form balanced clusters and prolong          | This algorithm is more suitable for static network but not for dynamic UWSN |
| CFC-PSO[17]      | Hybrid | to form energy-efficient clusters and select optimal CHs | Cluster count is fixed and initial CHs election in a random manner |
| FCM-PSO[21]      | Hybrid | to prevent early death of high loaded CHs      | The convergence speed is low when selecting CHs using PSO |

2) DETERMINATION OF THE OPTIMAL NUMBER OF THE CLUSTERS

In this section, we can use elbow method to determine the optimal number of the clusters, which can be realized by drawing the relation graph of \( k \) and cost function. The cost function can be expressed by the followed equation:

\[
F(k) = \sum_{j=1}^{k} \sum_{i=1}^{n} u_{ij}^{m} \|X_i - X_j\|^2 \quad (m \geq 0, \sum_{j=1}^{c} u_{ij} = 1) \tag{7}
\]

According to the equation, we can see the amount of fitness of the data can be evaluated by the cost function of the fuzzy clusters, including \( k \) cluster heads. Analyzing the cost function, we notice that \( F(k) \) decreases as the number of the cluster heads increases. In the beginning, the value decreases at a high speed, but after the optimal number \( c \) is determined, and as \( c \) increases, the value of the function, decreases at a lower speed [21].

3) DETERMINATION OF THE LOCATIONS OF THE INITIAL CLUSTER HEADS

After we obtain the membership degree matrix (\( k \times n \) matrix) of the clusters and the optimal number of the cluster heads \( c \), we can select \( c \) cluster heads from \( k \) initial cluster heads, and then convert the \( k \times n \) membership degree matrix to \( c \times n \) matrix. Now we can use FCM method to determine the locations of the cluster heads by iteration.

Thus far, through FCM method which regards the UWSN as fuzzy clusters, we has realized the initialization of the clusters. The reason for adopting FCM method is that the formation of the clusters is complex and the decency of the objects to the clusters is fuzzy in reality.

The following is the pseudocode that initializes cluster heads by the proposed method:

IV. SELECTION OF THE CLUSTER HEADS BY MFO METHOD

In section 3, we have obtained the initial cluster heads. The clustering algorithm aims at reducing the communication distances and energy consumption, so we should determine the optimal locations of the cluster heads. Therefore, selection of the cluster heads can be regarded as an optimization problem to reduce an objective function.

In this section, we will construct the objective function first, and then resolve the optimization problem through MFO method.

A. ENERGY MODEL IN UWSN

In UWSN, the energy consumption of the sensor nodes includes two parts: 1) Energy used by sensor nodes to sense the underwater environment, process the data packets, transmit the data packets to cluster heads, and receive data from cluster heads; 2) Energy used by the nodes that are cluster heads to receive the data packets from sensor nodes, and transmit and aggregate the data to the Base Station.

We express the energy consumption of a single node as follows [22]:

\[
E_{MEM} = E_{elec} + T_b C H d_{CM} e^{a(f)d_{CM}} \tag{8}
\]

where \( E_{elec} \) represents the loss of internal circuits such as amplifiers, encoders, underwater acoustic transducers, etc. \( T_b \) represents the bit duration of the data, and \( C H d_{CM} \) represents the distance between the node and the cluster head.

Similarly, the energy consumption of the cluster head can be expressed as the following formula:

\[
E_{CH} = E_{elec} + k(E_{elec} + T_b C H d_{CH} e^{a(f)d_{CH}}) \tag{9}
\]

where \( k \) represents the fusion ratio of data in the cluster head, and \( d_{CH} \) represents the distance between the cluster head and the Base Station.

In addition, \( a(f) \) is the absorption coefficient, and it can be calculated by using follow formula:

\[
a(f) = 0.11\frac{10^{-3}(f^2)}{1 + f^2} + 44*(10^{-3})\frac{f^2}{(4100 + f^2)} + 2.75*10^{-7}f^2*3*10^{-6} \tag{10}
\]

\( H \) is depth of the sensor nodes, and \( C = 2\pi*0.67*10^{-9.5} \).
Proposed Method to Initialize the Clusters

Function\(c,n\)

\[
k = \sqrt[2]{n/2}
\]

\[
\gamma = \left(\frac{E_r}{E}\right)^{-1} \parallel X_i - X_j \parallel
\]

for \(i = 1: n\)

for \(j = 1: k\)

\[n_k = \text{argmax} (\gamma)\]

end

end

\[\text{REPEAT}\]

\[
u_{ij} = \left[ \sum_{c=1}^{k} \left( \frac{\parallel X_i - V_j \parallel^2}{\parallel X_i - V_{ch} \parallel^2} \right) \right]^{-1}
\]

\[
V_j = \frac{\sum_{i=1}^{n} (u_{ij})^m X_i}{\sum_{i=1}^{n} (u_{ij})^m}
\]

\[\text{until} \parallel V_a - V_b \parallel < \varepsilon\]

end

\[\text{end}\]

\[F(k) = \sum_{c=1}^{c} \sum_{i=1}^{n} u_{ij} \parallel X_i - X_j \parallel^2 (m \geq 0, \sum_{j=1}^{c} u_{ij} = 1)\]

\[\text{draw the relation graph of} \ k \ \text{and cost function} \ F(k)\]

\[\text{determine the optimal number} \ c \ \text{of the clusters}\]

for \(i = 1: c\)

for \(j = 1: k\)

\[n_c = \text{argmax} (\gamma)\]

end

end

\[\text{REPEAT}\]

\[
u_{ij} = \left[ \sum_{c=1}^{C} \left( \frac{\parallel X_i - V_j \parallel^2}{\parallel X_i - V_{ch} \parallel^2} \right) \right]^{-1}
\]

\[
V_j = \frac{\sum_{i=1}^{n} (u_{ij})^m X_i}{\sum_{i=1}^{n} (u_{ij})^m}
\]

\[\text{until} \parallel V_a - V_b \parallel < \varepsilon\]

end

end

be affected by the communication distances. In fact, the selection of the cluster heads is an optimization problem, and the purpose of the optimization is to reduce the constructing the objective function.

1). The distance between a sensor node and a cluster head.

In FCM method, \(n\) nodes have been divided into \(m\) clusters, and the number of nodes in each cluster varies. So, we can express the distance between a sensor node and a cluster head of each cluster as follows:

\[
dist_{CM} = \sum_{j=1}^{N_j} d(CM_{ij}, CH_j) \quad (j = 1, 2, \ldots, c) \quad (11)
\]

where \(CM_{ij}\) is a member node in cluster \(j\), \(CH_j\) is a cluster head in cluster \(j\), and in each cluster, there are \(N_j\) nodes.

2). The distance between a cluster head and the base station

\[
dist_{CB} = d(CH_j, BS) \quad (12)
\]

where BS is the location of the base station.

3). Energy consumption of UWSN

The energy consumption of the networks for transmitting the data packets can be expressed as follows:

\[
E_{To} = \sum_{i=1}^{c} (E_{CH}^{i} + \sum_{j=1}^{N_i} E_{MEM}^{ij}) \quad (13)
\]

Through this formula, we can calculate the total energy consumption of the network, which divides \(n\) sensor nodes into \(c\) clusters each with \(N_i\) member nodes.

Based on the above 3 factors, we construct the objective function. In addition, considering there are differences among the scales of the 3 factors, we use the normalized function to deal with the differences. In this paper, we can use Sigmoid function to transform the 3 factors into \([0,1]\). Sigmoid function usually can be expressed as follows:

\[
S(x) = \frac{1}{1 + e^{-x}} \quad (14)
\]

Through the above, we can construct the objective function as follows:

\[
F_{obj} = W_1 dist_{CM} + W_2 dist_{CB} + W_3 E_{To} \quad (15)
\]

where \(W_1, W_2\) and \(W_3\) is the weight coefficient which can be adjusted to determine the priorities of the 3 factors. Simultaneously, \(W_1 + W_2 + W_3 = 1\).

C. SELECTION OF THE CLUSTER HEADS BY MFO

MFO algorithm is a nature-inspired algorithm which aims to find the best solutions for a particular problem. Moths can fly in night using a mechanism called transverse orientation for navigation. A moth flies by a fixed angle with respect to the moon, therefore, it can travel a long distance in a straight path. In fact, moths usually fly spirally around the artificial lights. This is because the artificial lights are not far enough like the moon. The moths try to keep the same angle with the lights, and they will maintain flying in a spiral path. Based on the motion model of the moths, MFO algorithm has been proposed. The particulars of the MFO process are described as follows [23]–[25]:

In MFO algorithm, the candidate solutions can be assumed to be moths, and the variables of the objective function are the locations of the moths in the searching spaces. The set of moths can be expressed as follows:

\[
M = \begin{bmatrix}
    m_{1,1} & m_{1,2} & m_{1,3} & \cdots & m_{1,d} \\
    m_{2,1} & m_{2,2} & m_{2,3} & \cdots & m_{2,d} \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    m_{n,1} & m_{n,2} & m_{n,3} & \cdots & m_{n,d}
\end{bmatrix}
\quad (16)
\]
where \( n \) is the number of the moths and \( d \) is the dimension of the variables. In addition, we can set an array to store the corresponding fitness values as follows:

\[
OM = \begin{bmatrix}
OM_1 \\
OM_2 \\
\vdots \\
OM_n
\end{bmatrix}
\]  

(17)

In this formula, the fitness value is the return value of the objective function for each moth. For example, we can substitute the position vector of one moth in the matrix \( M \) into the objective function, and then the return value of the objective function is the fitness value and will be assigned to the corresponding moth in \( OM \).

The flames are other components in MFO algorithm, and can be expressed as the followed matrix:

\[
F = \begin{bmatrix}
f_{1,1} & f_{1,2} & f_{1,3} & \ldots & f_{1,d} \\
f_{2,1} & f_{2,2} & f_{2,3} & \ldots & f_{2,d} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
f_{n,1} & f_{n,2} & f_{n,3} & \ldots & f_{n,d}
\end{bmatrix}
\]  

(18)

It should be noted that the flame matrix is similar to moth matrix, and there is an array to store the corresponding fitness values for the flames. The flame array can be expressed as follows:

\[
OF = \begin{bmatrix}
OF_1 \\
OF_2 \\
\vdots \\
OF_n
\end{bmatrix}
\]  

(19)

The moths are the search agents moving around the searching space, and the flames are the best positions obtained by moths at the present time. In this algorithm, moths can move around the flames and update them if they find better solutions. Therefore, moths will not lose their best solutions.

MFO algorithm should approximate the global optimal of the optimization problem and can be expressed as follows:

\[
MFO = (I, P, T)
\]  

(20)

\( I[M, OM] \) is a function which denotes a random initial population of the moths and the corresponding fitness values.

\[
M(i, j) = (ub(i) - lb(i))^\text{rand}() + lb(i)
\]  

(21)

\( ub \) and \( lb \) are upper and lower bounds of the variables, respectively.

\( P \) function moves the moths around the searching space, receives the values of \( M \) and returns the updated values of \( M \). \( M \) function can be expressed as follows:

\[
M_i = D_i^* e^{bt} \cos(2\pi t) + F_j
\]  

(22)

where \( M_i \) denotes the \( i \)-th moth and \( F_j \) denotes the \( j \)-th flame. \( D_i \) denotes the distance between \( M_i \) and \( F_j \) and can be expressed as:

\[
D_i = |F_j - M_i|
\]  

(23)

In addition, \( b \) is a constant defining the shape of the logarithmic spiral, \( t \) is a random number in \([-1, 1]\).

If the termination criterion is satisfied, the \( T \) function returns true. Otherwise, the \( T \) function returns false. The exit criteria is \( \|OF_{t+1} - OF_i\| < \varepsilon \), \( t \) is the number of iterations. If all the values in \( OF \) satisfy the exit criteria, the iteration will be stopped. After the initialization of the moths, the \( P \) function will iteratively run until the \( T \) function returns true.

In the proposed algorithm, selecting the cluster heads through MFO, we can initial \( S \) moths. Each moth can be denoted as

\[
MT_i = \{x_{i,1}, y_{i,1}, z_{i,1}; x_{i,2}, y_{i,2}, z_{i,2}; \ldots; x_{i,c}, y_{i,c}, z_{i,c}\}
\]  

(24)

where \( c \) is the optimal number of the cluster heads obtained from the above.

The following is the pseudocode that selects the cluster heads by the proposed algorithm:

V. SIMULATION AND RESULT

The simulation is carried out by MATLAB. In the experiments, it is considered that the nodes are dead when the energy of them is 0. In addition, when the nodes are dead, the performance of the network will be decreased. The more the nodes become dead, the more possibly the network black hole occurs. The efficiency of the proposed clustering algorithm based on FCM and MFO method which is defined as FCMMFO has been proved by the simulation.

In the experiments, we obtain the optimal number of the clusters, the locations of the cluster heads, and the clusters of the network by FCMMFO. Moreover, we compare the FCMMFO with FCMPSO, EEFCM-DE, LEACH algorithm in light of life cycle [15], [21], [26].

In the experiments, 100 sensor nodes are deployed in the area of 100m*100m*100m, the Base station is in (50,50,50), the size of data packet is 200bit, the data transmission rate is 1Kbps, the initial energy of each node is 0.5J, and the energy consumption of the circuit in each node is 50nJ/bit.

A. DETERMINATION OF THE NUMBER OF THE CLUSTER HEADS

Experiments are carried out by the algorithm in section 3 for 100 times, and the relation between the number of the cluster heads and the value of the cost function obtained is showed in Fig. 1. As shown in Fig. 1, when the number of the cluster heads is 3, the degree of distortion (y label) has been greatly improved. Based on the elbow method, we should select 3 as the optimal number of the cluster heads.

B. INITIALIZATION OF THE CLUSTERS

In the experiments, we have determined the optimal number of the cluster heads. Fig. 2 shows the change of the objective function. It shows that the function begins to converge when the iteration time is 50. In Fig. 3, the nodes in the networks have been divided into 3 clusters initially through FCM in section 2. Each cluster is displayed by different colored symbols.
Proposed Method to Select the Cluster Heads

\[
\text{Flame\_number} = \text{round}(N - \text{Iteration} \times (N - 1) / \text{Max\_iteration}) \quad \text{(where } N=S, \text{ which is the number of the moths)}
\]

Initial the moths:

\[
M(i, j) = (ub(i) - lb(i)) \times \text{rand()} + lb(i)
\]

\[
dist_{CM} = \sum_{i=1}^{N_i} d(CM_{ij}, M(i, j)) \quad (j = 1, 2, \ldots, c)
\]

\[
dist_{CB} = d(M(i, j), BS)
\]

\[
E_{To} = \sum_{i=1}^{c} (E_{M(i, j)} + \sum_{j=1}^{N_i} E_{MEM})
\]

Fitness(1,i) = objective function(\( dist_{CM}, dist_{CB}, E_{To} \))

If iteration = 1

\[
\text{[Fitness\_sorted I]} = \text{sort(Fitness)}
\]

\[
\text{sorted\_population} = M(1,:) ;
\]

\[
\text{best\_flames} = \text{sorted\_population} ;
\]

\[
\text{best\_flame\_fitness} = \text{fitness\_sorted} ;
\]

else

\[
\text{double\_population} = [\text{previous\_population}; \text{best\_flames}] ;
\]

\[
\text{double\_fitness} = [\text{previous\_fitness}; \text{best\_flame\_fitness}] ;
\]

\[
[\text{double\_fitness\_sorted I}] = \text{sort(double\_fitness)} ;
\]

\[
\text{double\_sorted\_population} = \text{double\_population}(1,:) ;
\]

\[
\text{fitness\_sorted} = \text{double\_fitness\_sorted}(1:N) ;
\]

\[
\text{sorted\_population} = \text{double\_sorted\_population}(1:N,:) ;
\]

\[
\text{best\_flames} = \text{sorted\_population} ;
\]

\[
\text{best\_flame\_fitness} = \text{fitness\_sorted} ;
\]

\[
\text{previous\_population} = M ;
\]

\[
\text{previous\_fitness} = \text{Moth\_fitness} ;
\]

\[
a = -1 + \text{Iteration} \times (-1) / \text{Max\_Iteration} ;
\]

for i = 1:size(M,1)

for j = 1:size(M,2)

if i < Flame_no

\[
distance\_to\_flame = \text{abs(sorted\_population(i,j)-M(i,j))} ;
\]

\[
b = 1 ;
\]

\[
t = (a-1) \times \text{rand()} + 1 ;
\]

\[
M(i,j) = \text{distance\_to\_flame} \times \exp(b \times \cos(t \times 2 \times \pi)) + \text{sorted\_population(i,j)} ;
\]

end

if i > Flame_no

\[
distance\_to\_flame = \text{abs(sorted\_population(i,j)-M(i,j))} ;
\]

\[
b = 1 ;
\]

\[
t = (a-1) \times \text{rand()} + 1 ;
\]

\[
M(i,j) = \text{distance\_to\_flame} \times \exp(b \times \cos(t \times 2 \times \pi)) + \text{sorted\_population(Flame\_no,j)} ;
\]

end

end

iteration = iteration + 1 ;

end

C. SELECTION OF THE CLUSTER HEADS USING MFO

In the previous section, we obtain the clusters initially. Based on the initial clusters, we simulate and obtain the locations of the cluster heads. Fig. 4, Fig. 5 and Fig. 6 show the cluster heads obtained by MFO, PSO and EEFCM-DE, respectively. Such cluster heads are denoted by the symbol “o” and in the same color of their corresponding clusters. Fig. 7 shows the convergence curve of the fitness function, which shows that clustering algorithm using MFO has higher convergent speed.

D. LIFE CYCLE OF THE NETWORK

In the experiments, we simulate the energy consumption of LEACH, FCMPSO, EEFCM-DE and proposed FCMMFO method. Fig. 8, Fig. 9 and Fig. 10 depict the number of the alive nodes and the rounds. In Fig. 7, the weight coefficients are set as followed: W1=0.5, W2=0.1 and W3=0.4. At the same time, in Fig. 8, the weight coefficients are as followed:
W1=0.4, W2=0.1 and W3=0.5. In Fig9, the weight coefficients are as followed: W1=0.3, W2=0.1 and W3=0.6. The initial energy and the deployment of the nodes in Fig. 8, Fig. 9 and Fig. 10 are the same.

Fig. 8, Fig. 9 and Fig. 10 show that both FCMMFO, EEFCM-DE and FCMPSO are superior to LEACH in prolonging the lifetime of the network, under the condition that the initial simulation parameters, except for the weight coefficients, are the same. As to energy consumption, the figures show FCMMFO-based network lifetime increases, and the FCMMFO actually consumes less energy, due to its less computing time with higher convergent speed.

From Fig. 8, Fig. 9 and Fig. 10, it can be concluded as followed: in LEACH method, all nodes become dead within about 1000 rounds with the first dead node appearing nearly in the 160 round, and there are new dead nodes almost
in each round; while in both FCMMFO,EEFCM-DE and FCMPSO methods, all nodes become dead within about 1600 rounds with the first dead node appearing approximately in the 1200th round(Fig. 8),1400th round(Fig. 9) and 1600th (Fig. 10). The most important reason for the above conclusion is that the division of the clusters and the selection of the cluster heads help prevent the early death of the nodes by reducing the energy consumption of the communications between the sensor nodes and balancing the consumption of each node in UWSN effectively. In addition, in Fig. 8, Fig. 9 and Fig. 10, the weight coefficients influence the performance of the network, thus, the optimal performance can be found out by adjusting the weight coefficients.

To present the performance of FCMMFO-network in different scenarios, the results for average energy consumed per round for the network is evaluated as Table 2 above. In the simulation, the number of nodes varies from 50 to 200, and the simulation conditions is: initial energy of each node is 0.5J, the weight coefficients of fitness function are as followed: W1=0.3, W2=0.1 and W3=0.6. Table 2 shows that the average energy consumed per round for the network is significantly small in different scenarios. The results show that FCMMFO is energy-efficient in all the scenarios.

**TABLE 2. Average energy consumed per round.**

| Number of nodes | Average energy consumed per round(units:J) |
|-----------------|------------------------------------------|
| 50              | 0.0002116                                |
| 100             | 0.0003021                                |
| 150             | 0.0004013                                |
| 200             | 0.0005176                                |

**VI. CONCLUSION**

This paper focuses on the clustering algorithm of UWSN, of which sensor nodes are deployed randomly in a 3D underwater area. In order to obtain an energy-efficient clustering protocol, according to the FC and MFO algorithm, we propose a hybrid algorithm to cluster and select the CHs. The proposed algorithm initially divides the network into clusters through FC, which can obtain the membership based on the distances between any 2 nodes. In addition, we use elbow method to determine the optimal number of the clusters. Further, through MFO method, we construct the objective function based on 3 factors: the distance between a node and a cluster head, the distance between a node and Base station, and the energy consumption of the network. Then, we find out the best solutions, which represent the optimal locations of the cluster heads.

Simulation results show that the proposed algorithm obtains better performance compared with LEACH, because it effectively reduces the energy consumption of the entire UWSN and balance the energy consumption of each node. Compared with FCMPSO,EEFCM-DE algorithm, the proposed algorithm has higher convergent speed, thus it will reduce the energy consumption of the network. In addition, the weight coefficients in objective function influence the performance of UWSN. The larger the value of W3 in (15) is, the longer the lifetime of the network is, because W3 is the weight coefficient of the energy consumption of the network.

In summary, on the one hand, the proposed algorithm is a dynamic protocol suitable for the dynamic UWSN, as it selects the CHs based on the residual energy and the distances between the sensor nodes. On the other hand, the proposed algorithm provides an effective method to deploy mobile nodes or elect the nodes as CHs for a large scale UWSN. However, there are some open issues, such as the time interval for the selection of the CHs, that still need to be further discussed.

**REFERENCES**

[1] J.-H. Cui, J. Kong, M. Gerla, S. Zhou, “The challenges of building mobile underwater wireless networks for aquatic applications,” *IEEE Netw.*, vol. 20, no. 3, pp. 12–18, May/Jun. 2006.

[2] L. Lanbo, Z. Shengli, and C. Jun-Hong, “Prospects and problems of wireless communication for underwater sensor networks,” *Wireless Commun. Mobile Comput.*, vol. 8, no. 8, pp. 977–994, 2008.

[3] I. F. Akyildiz, D. Pomplili, and T. Melodia, “Underwater acoustic sensor networks: Research challenges,” *Ad Hoc Netw.*, vol. 3, no. 3, pp. 257–279, 2005.

[4] D. N. Sandeep and V. Kumar, “Review on clustering, coverage and connectivity in underwater wireless sensor networks: A communication techniques perspective,” *IEEE Access*, vol. 5, pp. 11176–11199, 2017.

[5] S. Yadav and V. Kumar, “Optimal clustering in underwater wireless sensor networks: Acoustic, EM and FSO communication compliant technique,” *IEEE Access*, vol. 5, pp. 12761–12776, 2007.

[6] 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.

[7] B. Singh and D. Lobiyal, “A novel energy-aware cluster head selection based on particle swarm optimization for wireless sensor networks,” *Hum.-Centric Comput. Inf. Sci.*, vol. 2, no. 1, p. 13, 2012.

[8] P. C. S. 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, pp. 2005–2020, Apr. 2017.

[9] H. Zhang, S. L. Wang, and H. X. Sun, “A robust clustering architecture for medium scale Underwater acoustic sensor networks,” (in Chinese), *J. Xiamen Univ. Nat. Sci.*, vol. 2014, Art. no. 647281.

[10] 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, pp. 9671, 2019.

[11] X. Zhao, Z. Wei, Y. Cong, and B. Yin, “A balances energy consump- tion clustering routing protocol for a wireless sensor network,” in *Proc. IEEE 4th Int. Technol. Mechatronics Eng. Conf. (IOTEC)*, Dec. 2018, pp. 521–527.
W. Fei et al.: Energy-Efficient Clustering Algorithm in UWSNs Based on FCM and MFO Method

[15] R. Sharma, V. Vashisht, and U. Singh, “EEFCM-DE: Energy-efficient clustering based on fuzzy c means and differential evolution algorithm in WSNs,” IET Commun., vol. 13, no. 8, pp. 996–1007, May 2019.

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

[17] A. Javadpour, N. Adelpour, G. Wang, and T. Peng, “Combining fuzzy clustering and PSO algorithms to optimize energy consumption in WSN networks,” in Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Adv. Trusted Comput., Scalable Comput. Commun., Cloud Big Data Comput., Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Oct. 2018, pp. 1371–1377.

[18] S. Potthuri, T. Shankar, and A. Rajesh, “Lifetime improvement in wireless sensor networks using hybrid differential evolution and simulated annealing (DESA),” Ain Shams Eng. J., vol. 9, no. 4, pp. 655–663, Dec. 2018.

[19] A. Shokrollahi and B. Mazloom-Nezhad Maybodi, “An energy-efficient clustering algorithm using fuzzy C-Means and genetic fuzzy system for wireless sensor network,” J. Circuits, Syst. Comput., vol. 26, no. 01, Jan. 2017, Art. no. 1750004.

[20] F. Wang, L. Wang, Y. Han, B. Liu, J. Wang, and X. Su, “A study on the clustering technology of underwater isomorphic sensor networks based on energy balance,” Sensors, vol. 14, no. 7, pp. 12523–12532, 2014.

[21] V. Krishnaswamy and S. S. Manvi, “Fuzzy and PSO based clustering scheme in underwater acoustic sensor networks using energy and distance parameters,” Wireless Pers. Commun., vol. 108, no. 3, pp. 1529–1546, Oct. 2019.

[22] N. Goyal, M. Dave, and A. K. Verma, “Energy efficient architecture for intra and inter cluster communication for underwater wireless sensor networks,” Wireless Pers. Commun., vol. 89, no. 2, pp. 687–707, Jul. 2016.

[23] S. Mirjalili, “Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm,” Knowl.-Based Syst., vol. 89, pp. 228–249, Nov. 2015.

[24] R. Barham, A. Sharieh, and A. Sleit, “Multi-moth flame optimization for solving the link prediction problem in complex networks,” Evol. Intell., vol. 12, no. 4, pp. 563–591, Dec. 2019.

[25] S. J. Nanda, “Multi-objective moth flame optimization,” in Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI), Sep. 2016, pp. 2470–2476.

[26] X. W. Ma and X. Yu, “Improvement on LEACH protocol of wireless sensor network,” Appl. Mech. Mater., vol. 347–350, pp. 1738–1742, Aug. 2013.

WANG FEI was born in Taiyuan, Shanxi, China, in 1987. He received the M.S.E. and Ph.D. degrees in signal and information processing from the North University of China, Taiyuan, in 2012 and 2014, respectively. He is currently a Lecturer at the School of Computer and Information Technology, Shanxi University, Taiyuan. His current research interests include optimization in performance, routing, and energy consumption of UWSN.

WANG JIANJUN was born in Jiangsu, China. He received the M.S. degree in electronics and communication engineering from the North University of China, in 2014. He is currently pursuing the Ph.D. degree in communication and information system with the Nanjing University of Posts and Telecommunications. He is also an Engineer at Grain and Oils Information Center of Jiangsu Province. His research interests include cyber security and informatization, big data and artificial intelligence applications, digital signal processing, and telecommunications.

BAI HEXIANG is currently an Associate Professor of computer science at the School of Computer and Information Technology, Shanxi University. His research interests include rough set-based granular computing.

LI DEYU is currently a Professor of computer science at the School of Computer and Information Technology, Shanxi University. His research interests include rough set-based granular computing.