Proc. of Int. Conf. on Computing Electrical and Electronics Engineering, CEEE

Enhanced Energy for Data Aggregation in Wireless Sensor Networks based on Clustering Techniques

Prof.S.Kokilavani[1] and Dr.N.Sathishkumar[2]
[1]Principal, Hindusthan Polytechnic College, Coimbatore-32, Tamil Nadu, India
Email: vani.msb@gmail.com
[2]Professor, Department of Electronics & Communication Engineering, Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India
Email: nsk20022002@gmail.com

***Abstract*—Enhanced energy for data aggregation in wireless sensor networks based on clustering techniques uses fuzzy-c clustering algorithm to perform energy efficient data aggregation of the wireless sensor network models. In sensor network models, always data aggregation is of prime importance due its dependency on numerous factors like topology of the network, link factors, and energy constraints and so on. Traditional techniques fail in energy efficient data aggregation because of the node's battery power and the degradation of network lifetime. Hence, in this research paper attempt is made to improve the data aggregation and network lifetime of the sensor network model employing a new enhanced optimization algorithm (EOA) based fuzzy c-means (FCM) clustering technique. The new enhanced optimization algorithm is developed considering the worst positions traversed by the population during the optimization process and hence moving towards attainment of better optimal cluster points. Fuzzy c-means clustering is designed in this paper by employing ensemble elbow technique based data aggregation process. The newly developed EOA algorithm is hybridized with the fuzzy c-means clustering technique so as to model a novel hybrid EOA-FCM technique for better data aggregation and improving the lifetime of the sensor network models. In this work, Bray-Curtis similarity index is employed to evaluate the similarity between the sensor nodes and that of the cluster nodes. Simulation is carried out for the considered network model employing the proposed EOA-FCM technique and the set performance metrics are evaluated to prove the superiority of the developed hybrid EOA-FCM technique in comparison with the existing approaches from literatures.**

***Index Terms*— Fuzzy c-means clustering, optimization technique, Hybrid clustering algorithm, Fuzzy optimal algorithm, and Data aggregation.**

I. INTRODUCTION

A wireless sensor network (WSN) is built with sensor devices to gather and organize the environmental data and here sensor nodes are deployed to coordinate with each other to execute the assigned task. Since wireless sensor networks are employed in numerous real time applications, the energy of sensor nodes is a promising factor wherein the node senses huge volume of information and transfers voluminous of data. Existence of minimal energy level results in increasing the inactive nodes and then the entire network becomes failure in maintaining the energy resources. Considering this, the major issues in sensor network models are the energy efficient data

Grenze ID: 02.CEEE.2022.7.503

aggregation and the improvement of network lifetime. In this aspect, clustering algorithms has always been a solution to solve these issues and over the years, numerous techniques have been developed to perform data aggregation and also to increase the lifetime of the network. But each of the approaches possess their own merits and demerits, hence is the development of this novel hybrid EOA based FCM technique to carry out effective and better data aggregation of the network models. Figure 1 presents a sensor network model with data aggregator between the sensor nodes and the base station. Figure 2 provides the clusters formed among the nodes and each of the aggregator nodes pertaining to the respective cluster communicating to the base station.

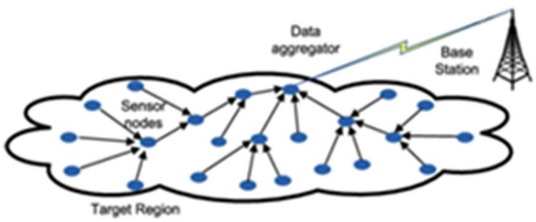


Figure 1. Data Aggregation in Sensor Network Model

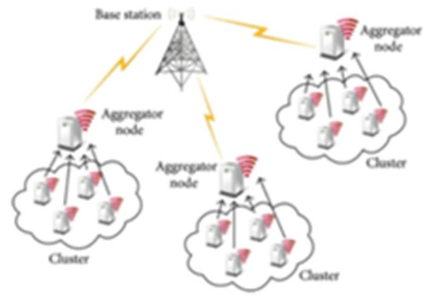


Figure 2. Clustering Mechanisms for Data Aggregation

In connection with the data centric routing, data aggregation play a major role for solving overlap problems and the data from the multiple sensor nodes get aggregated like they are of the same attribute on reaching the similar routing node when it returns to the sink. Energy efficiency during data aggregation, network lifetime and data integrity become key factors when the sensor networks are employed in an aggressive environment. Figure 3 presents the basic flow of data aggregation process. The aggregated data employing the techniques developed in finally transferred to the sink node by choosing an effective path.

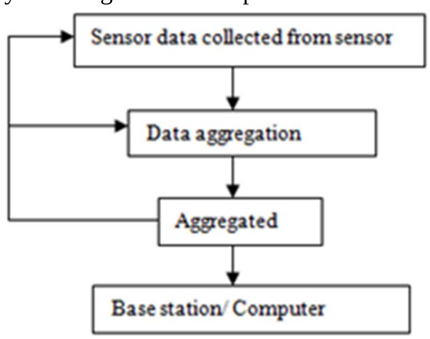


Figure 3. Flow Process of Data Aggregation

The operating principle of data aggregation is by selecting the nodes and making them divided into clusters. The clusters are selected in a manner that they satisfy the conditions of the parameters and as well satisfy the number of nodes pertaining to that cluster. Pertaining to the clusters formed, in each cluster, a cluster head (CH) is chosen. These cluster heads will be responsible for data collection within each cluster and so as to transfer the data to the nearby cluster head for exchange of required information. Based on the cost criterion, cluster head gets altered and each of the cluster nodes satisfying the cost criterion shall become cluster heads and all the nodes will be provided with an opportunity to act as cluster heads. At the time of data aggregation process, it addresses numerous user queries and is transformed to low level schemes with a query processor. Finally all the data will be transferred to the base station for further use.

In this juncture, numerous methods have been adopted [1-7] over the years to perform effective and superior energy efficient data aggregation and thereby to improve the lifetime of the sensor network models, but there has always been a lacuna in transfer of data, increased energy rate and always there is a requirement for development of new techniques for efficient and effective data aggregation and network lifetime improvement of sensor network models. The growth of machine learning techniques has paved the way for developing more clustering based techniques and hence this work focuses on developing a new optimized fuzzy c-means clustering based energy efficient data aggregation model for the considered sensor network model. Machine learning based fuzzy algorithms with their rule inference mechanism tends to find solutions to various complex problems and due to which in this research paper it is intended to model an optimized fuzzy c-means clustering model to execute the task of data aggregation in sensor network models.

II. Related Works

The development of machine learning based neural computing, fuzzy computing and evolutionary computing has paved their way in several applications and always there has been a focus for developing better data aggregation model for the sensor networks. Considering the merits of these growing machine learning based computing methods, in this research paper attempt is taken to model a new EOA-FCM technique for effective data aggregation. This section details the review made on the previous techniques developed for the considered application.

A Multi-Mobile Agent Itinerary Planning-Based Energy and Fault Aware data aggregation (MAEF) was introduced in a work by Fissaoui et al [1]. But, here the network lifetime was not improved using MAEF and this was the limitation of the developed model. A Sparsest Random Sampling method for cluster-based compressive data gathering (SRS-CCDG) was designed with lesser energy cost [2].The designed method does not show proving results with respect to the data gathering accuracy.

An Energy-aware Compressive sensing based Data Aggregation (ECDA) model was introduced to overcome the problem of network lifetime and reduce the energy consumption [3]. However, the data aggregation performance was significantly minimal. A sparsity feedback-based compressive data aggregation method was designed for balancing the energy between the nodes [4]. Mixed-integer linear programming (MIP) model was developed for energy efficient data collection. But in this case also, the data collection time was not minimized [5]. To conserve node energy and improve the network lifetime, a dynamic mobile agent-based data collection method was presented. But this technique failed to enhance the accuracy of data aggregation [6].

A low delay and high-throughput opportunistic data collection scheme were introduced to improve the DP transmission to sink [7]. The designed scheme was not proven for the energy-aware data collection. A clustering method was introduced for energy efficient data gathering and enhancing the lifetime of the network. But employing this technique the clustering error was not minimized [8].A Distributed Data Gathering Approach (DDGA) was developed for solving the data aggregation problem with a mobile sink [9].The approach decreases the energy utilization but the accurate data gathering was not achieved with this developed model. A prediction model-based data collection was developed using the cluster heads for improving the collected data's accuracy with a minimum predefined error. Though the model improves the lifetime of the node, the data collection time was not minimized with this proposed approach [10]. A new energy-efficient data collection approach was introduced with spatial-temporal correlation for achieving the reasonable accuracy. But here also there was a limitation in the performance of network lifetime [11]. A probabilistic clustering algorithmwas introduced to perform data aggregation of the sensor network models [12].

To enhance data aggregation performance and lessen the energy consumption, a neural network model was also designed [13]. The time taken for efficient data gathering remains unsolved. A reinforcement learning based clustering algorithm (RLBCA) was developed to find cluster heads for collecting the data and to send it to the sink node. But in this case, the clustering error was observed to be higher [14].

A spawn multi-mobile agent itinerary planning (SMIP) method was developed to enhance the data gathering processes with minimal energy utilization and time. The performance of network lifetime was not increased employing the proposed approach [15].A Multi-Strip Data Gathering (MSDG) method was designed to lessen the energy utilization of sensor nodes [16]. But in this case also, the time complexity of data gathering was not minimized. A resilient data aggregation scheme was developed with the spatio temporal correlation for the sensor network model [17].

To gather the data from cluster heads, a novel energy–aware and density-based clustering algorithm was designed [18]. A novel itinerary planning algorithm grouping the mobile node based on the density was developed [19]. Using this developed approach, an effective data gathering was not performed with the cluster heads. For enhancing the data gathering efficiency and stability of energy consumption, a type of data gathering technique with mobile sink was designed [20]. In this case also, the data gathering time was not minimized.

Considering the review presented above, it is well noted that numerous techniques have been developed over the past years for effective data aggregation of sensor network models. The review made has established that limitations were noted during the implementation of the developed techniques and henceforth this paper is designed to bring out the advantages of machine learning based fuzzy computing method for the data aggregation of the sensor network models. In the developed optimized fuzzy c-means clustering, the novelty lies in its optimal cluster identification employing newly developed enhanced optimization algorithm and the applicability of ensemble elbow method and Bray-Curtis similarity index is employed to formulate a similarity metric. Based on this, the main contributions made in this paper includes,

- To develop a new enhanced moth-flame optimization algorithm to find the optimal clusters of the sensor nodes.
- To model the optimized fuzzy c-means clustering based on new EOA and ensemble elbow method for obtaining a better data aggregation model.
- To incorporate elbow technique to determine suitable clusters and present it to the optimized fuzzy c-means clustering algorithm. Here, Bray-Curtis Similarity Index is employed to identify the higher residual energy nodes and thereby they are employed for data aggregation.
- To minimize the data aggregation time and increasing the network lifetime employing the proposed hybrid EOA-FCM approach.

III. PROPOSED NEW HYBRID EOA FCM TECHNIQUE

In this section a new hybrid optimized fuzzy c-means clustering model is developed combining the best characteristics of the enhanced moth-flame optimization algorithm and the elbow technique based fuzzy c-means clustering algorithm. The key objective of the new modeled hybrid EOA-FCM technique is that the optimal clusters will be identified using the enhanced optimization algorithm and then based on the cluster heads and sensor node activations, effective data aggregation meeting the energy requirements will be carried out for the considered sensor network model.

A. *Proposed ensemble elbow based Fuzzy C-means clustering*

Fuzzy C-means algorithm is in cluster analysis, pattern recognition, image processing, and so on. In sensor network models, FCM groups sensor nodes to appropriate cluster based on membership degree and geographical locations. The nodes with greater residual energy in sensor networks are used to gather data and send the information. In this paper, the nodes scrutinize the energy level using FCM algorithm for improving the network life time. By applying the FCM, each sensor node energy level is estimated to become the probability of cluster head. After scrutinizing the nodes energy level, the node with greater energy level is chosen as the cluster head for each cluster. The FCM combines various clustering factors for the selection of cluster heads. Here, data transmission from each sensor Employing the new ensemble elbow based FCM approach, selecting the nodes with higher residual energy as cluster heads increases the performance of entire system and increase the networks life time. Fuzzy decision rules are developed here to select intelligent cluster heads. The selection of cluster heads significantly influences the network performance. Hence, the newly modeled FCM elects energy efficient cluster head among the number of nodes in order to gather the sensed data from the neighboring nodes. Then, it transmits the aggregated data to the sink node for further operations.

B. Developed ensemble elbow based fuzzy C-means clustering for data aggregation

In sensor networks, each sensor nodes are deployed with an equal energy level. Energy of node is calculated as product of power and time which is calculated using,

$$E = [power * time] \tag{1}$$

In equation (1), 'E' indicates energy of sensor nodes. Due to sensing nature of node, the energy level gets degraded. Therefore, the residual energy of sensor node is estimated. The residual energy is remaining node energy after sensing the data. Residual energy is evaluated using,

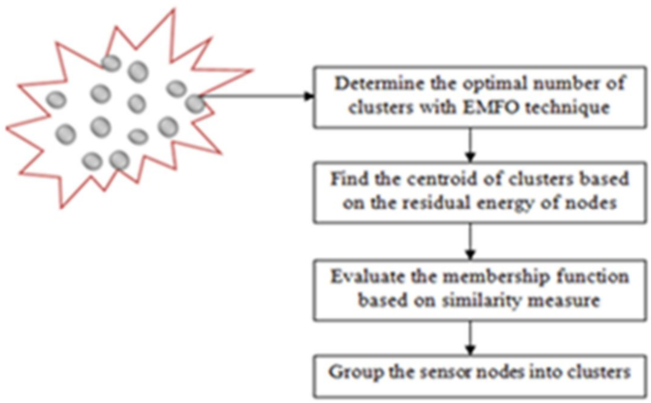
$$E_r = \{E_{total} - E_{consumed}\} \tag{2}$$

In equation (2), 'E_r' indicates residual energy of sensor, E_{total} is the total energy of node, $E_{consumed}$ indicates the consumed energy of node. Combining ensemble elbow method and fuzzy c-means clustering algorithm, the sensor nodes are grouped.

The proposed new EOA technique in previous section is used here to find optimal 'c' number of clusters and then the residual energy based clustering is done by using the fuzzy c-means algorithm. The conventional clustering techniques randomly initialize the number of clusters resulting in occurrence of errors after the clustering process. In order to overcome this limitation, the proposed technique introduces the elbow method based FCM to carry out effective grouping process. Figure 4 provides the flow process of elbow technique based FCM approach.

In figure 4, number of sensor nodes is disseminated in network. On applying enhanced moth-flame optimization technique, optimal 'c' number of clusters is selected for grouping the sensor nodes. The elbow criterion minimizes the summation of square error of clustering and on increasing the value of 'c', the error gets decreased. Therefore, it is clear that higher number of clusters results in error minimization. For cluster value to be 1, the mean square error is obtained using,

$$\omega = \sum_{i=1}^{n} \sum_{j=1}^{c} \left\| SN_i - C_j \right\|^2 \quad (3)$$



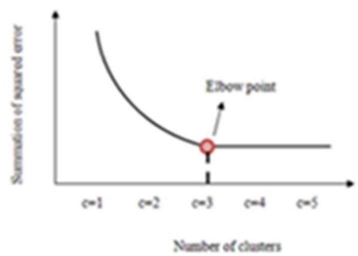


Figure 4. Flow Process of EOA based Elbow-FCM Technique   Figure 5. Cluster Selections EOS-FCM technique

In equation (3), 'ω' denotes the summation of the squared error, SN_i denotes the sensor nodes in the cluster, C_j is the 'j-th' cluster. On increasing the cluster value to be $c = 2$ and $c = 3$,then three clusters are formed and error is calculated. Similarly, the error gets decreased on increasing the number of clusters. At a particular point, the elbow point is met and the resultant cluster is the optimal value of clusters.

In Figure 5, the numbers of clusters are denoted with $c = 1, c = 2, c = 3, c = 4, c = 5$ and red color circle denotes the elbow point. As shown in figure 5, three optimal clusters (i.e. $c = 3$) are chosen for grouping the sensor nodes. Depends on their energy level, number of clusters are chosen for grouping the sensor nodes. Once the optimal cluster is selected with EOA approach, then centroid is defined for each cluster. Fuzzy membership is computed based on the similarity measure.

$$\beta_{ij} = \frac{1}{\sum_{i=1}^{n} \sum_{j=1}^{c} \left(\frac{\rho_{ij}}{\rho_{ic}}\right)^{\left(\frac{2}{n-1}\right)}} \quad (4)$$

In equation (4), 'β_{ij}' represents the fuzzy membership function which is used to identify the member of that particular cluster, ρ_{ij} denotes a similarity among i^{th} SNs and j^{th} clusters center, ρ_{ic} represents the similarity between the 'i^{th} SNs and c^{th} cluster centroid.

Similarity among sensor nodes and cluster centroid is calculated using Bray-Curtis Similarity Index.

$$\rho_{ij} = 1 - \frac{2\,(d_{ij})}{|SN_i| + |c_j|} \quad (5)$$

In equation (5),ρ_{ij} denotes a Bray-Curtis Similarity coefficient, d_{ij} denotes a mutual independence between the sensor nodes and cluster center, $|SN_i|$ and $|c_j|$ represents the cardinalities of the two sets (i.e. number of an element in each set i.e. SNs and cluster centers). The Bray–Curtis similarity coefficient is bounded between 0 and 1, where 1 means the sensor node is a favorable node to that cluster and 0 means the sensor node is the non-favorable node to that cluster. The process gets repeated until all the sensor nodes get clustered to the suitable clusters. For each iteration, the cluster centroid gets updated to group all the sensor nodes.

$$c_j(t) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{c} (\beta_{ij})^{\tau} * SN_i}{\sum_{i=1}^{n} \sum_{j=1}^{c} (\beta_{ij})^{\tau}} \quad (6)$$

In equation (6),$c_j\,(t)$ denotes an updated cluster centroid of the clusters, τ denotes a fuzzifier which determines the level of cluster fuzziness. β_{ij} indicate fuzzy membership function. Based on the updated value of cluster center and residual energy level, the entire sensor nodes are grouped into cluster. Figure 6 presents the flowchart that partitions the complete wireless network into different groups employing Bray-Curtis similarity. The existence of higher similarity among the energy level of the sensors and centroid possess higher probability to group sensor nodes into particular cluster. This process gets iterated until convergence is attained. Once the node clustering is completed, then the cluster head is selected for aggregating the data. Each group possesses one cluster head and its respective cluster members. The sensor node possessing higher residual energy is chosen as the cluster head. Cluster head receives the data from its respective member nodes and sends it to sink node. Figure 7 provides the data aggregation carried out based on cluster mechanism.

C. Proposed new EOA based Fuzzy c-means Alustering Algorithm

The hybrid enhanced moth-flame optimization algorithm based ensemble fuzzy c-means algorithm is developed in this section. Initially the population of sensor node is generated and enhanced moth-flame optimization technique is used to determine the optimal number of clusters and then the proposed new ensemble FCM approach in section 3.3 is used. With respect to the residual energy, sensor nodes get grouped into different clusters. The enhanced EOA technique with its updated flame position and moth position applied transverse orientation to find the optimal cluster. The fitness function with which EOA technique operates is the mean square error given in equation (6). After that, the elbow based FCM approach group the sensor nodes with high similarity. The centroid gets assigned to the cluster based on the residual energy of the node. Membership is evaluated that discovers member belonging to particular cluster. Similarity is computed to discover the favorable and non-favorable node of the cluster. Based on the similarity index, the favorable nodes gets grouped into clusters and cluster head is selected for coordinating sensor nodes within the cluster. Finally, sensor node passes the information to the cluster head and this cluster head aggregates the data from its member and sends it to sink node. The information is finally received at the sink node and this complete hybrid mechanism attains better energy efficient data aggregation with minimal time consumption. Table 1 provides the developed new hybrid EOA-FCM algorithm for data aggregation.

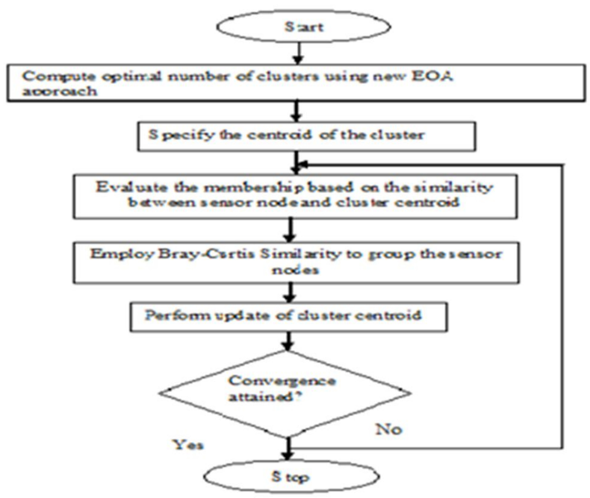


Figure 6 Flow Chart of the proposed technique

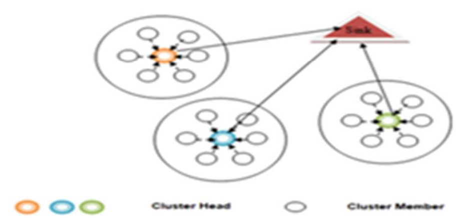


Figure 7 Data aggregation based on cluster mechanism

D. Simulation results and discussions

For enhancing the data aggregation and to minimize the energy consumption, the hybrid EOA based elbow fuzzy c-means clustering has been developed in this paper and simulations is being carried out by performing appropriate partitions to identify the cluster heads. The partitions are done based on the energy level of sensors and each group comprises of a cluster head and respective cluster members under them. This process of cluster head based data aggregation consumes minimal energy and the durability of the network increases.

E. Network model for realization of proposed optimized FCM approach

The network model is described in this section with the number of spatially distributed sensor nodes. Sensor network is organized into the directed graph $G_d = (v, e)$which comprises the set of vertex (v) and edge (e). The vertex (v) represents the 'n' sensor nodes $\mathrm{SN}_1, \mathrm{SN}_2, \mathrm{SN}_3, \ldots\ldots \mathrm{SN}_\mathrm{n}$distributed in the squared area $N * N$. In a graph, 'e' isan edge which represents the link i.e. connection between the SNs. Total network is divided into 'c' number of clusters $\mathrm{C}_1, \mathrm{C}_2, \mathrm{C}_3, \ldots \mathrm{C}_\mathrm{c}$based on node residual energy level($\mathrm{i.e}\ \mathrm{E}_\mathrm{r}$). Cluster head is chosen for every group. Sensor nodes gather the data $\mathrm{DP}_1, \mathrm{DP}_2, \mathrm{DP}_3, \ldots . \mathrm{DP}_\mathrm{n}$ from the environmental conditions and transmit it to the cluster head. Cluster head transmits the collected data to sink node (S_n) where it act as data aggregator.

Based on the specified network model, the proposed hybrid EOA-Elbow FCM technique is designed and modeled to carry out effective and efficient data aggregation.

F. Parameter settings for the proposed model

Simulation process for the proposed EAO-FCM technique and other existing methods for comparison [1, 2] are modeled and implemented in NS2.34 network simulator. Sensor nodes are disseminated in a square area of A^2 (1100 m * 1100 m) with random waypoint mobility model. Sensor nodes are moved in the speed of 0 to 20m/sec. Simulation time is set to be 300 sec. For energy efficient data aggregation, the Dynamic Source Routing (DSR) protocol is used in the simulation setup. Table 2 provides the parameter settings of the network model and Table 3 presents the simulation parameters for the proposed EOA-FCM technique.

G. Results and analysis

The results of the proposed EOA-FCM technique and that of the existing MAEF [1] and SRS-CCDG [2] are simulated and compared to prove the superiority of the developed approach. The parameters energy consumption, network lifetime, data aggregation accuracy and data aggregation time are evaluated using the proposed technique and the results are presented in the below given sub-sections.

H. Energy consumption

Energy consumption is defined as the amount of energy taken by the sensor node to aggregate data points. The energy consumption formula is evaluated using,

$$E_{con} = SN_n * E_{con} \ (single\ SN) \qquad (7)$$

In equation (7),E_{con} denotes energy consumption, SN_n represents the number of sensor nodes (SN). Energy consumption is measured in joule (J).To evaluate the energy consumption, the sensor nodes are considered in the range of 50-500. On consideration of 50sensor nodes to perform the simulation, the proposed EOA-FCM technique attains 28joules of energy consumption whereas the state-of-the-art works MAEF technique [1] and SRS-CCDG approach [2] consumes31 joules and 35 joules. Based on the above discussion, it is well clear that the developed EOA-FCM method consumes minimal energy than the existing works from previous literatures.

TABLE 1. PROPOSED NEW HYBRID EOA-FCM ALGORITHM FOR DATA AGGREGATION

Start
Input – Number of sensor nodes (SN_1, SN_2,…,SN_n) and number of data packets (DP_1, DP_2, …, DP_n)
Output – Residual energy, increased network life time
Initialize the population of sensor nodesM_pop_m
Invoke: EOA process
Generate M_pop_m(m = 1,2,…,n)
WHILE *(best_fitness!=present_fitness)*
Iteration=iteration+1
Compute fitness function as given in equation (14)
FF_{M_pop}=f(M_pop)
Update the flame position, FL_n (n=1,2,…,n)
IF *(iteration==1)*
$FL_{fitness}$=sort(FF_{M_pop})
FF_n=sort(M_pop_m)
ELSE
IF *sort(FF_{M_pop})<$FL_{fitness}$*
$FL_{fitness}$=sort(FF_{M_pop})
FF_n=sort(M_pop_m)
ENDIF
ENDIF
Update mean flame positions, $FL_{n1,n2}$ using equation (9)
$FL_{n1,n2} \in FL_{n1}, FL_{n2}$
Update moth_position using equation (11)
END WHILE
Return Best $M\_pop_m \rightarrow SN_i$
Invoke: Ensemble Elbow FCM algorithm
Evaluate energy E and residual energy E_r
Obtain the optimal number of clusters to be Best M_pop_m from EOA technique
FOR *each chosen clusters*

Define cluster centroid
Evaluate membership β_{ij} based on the similarity ρ_{ij}
Measure the similarity ρ_{ij}
***IF** (ρ_{ij} =1) THEN*
SN_i grouped into the suitable clusters
ELSE
SN_i becomes the non-favorable mode to that cluster
ENDIF
***FOR** each cluster c_j*
Select the cluster head (C_H) with higher residual energy
C_H collects data from its members
C_H sends data packet to sink node
ENDFOR
ENDFOR

Stop

Table 2 Simulation parameters for the network model

Simulation Parameters	Parametric values
Network simulator	NS2.34
Square area	1100 m×1100 m
Number of sensor nodes	50, 100, 150, 200, 250, 300, 350, 400, 450, 500
Mobility model	Random waypoint model
Speed of sensor nodes	0 – 20 m/sec
Simulation time	300 sec
Protocol	DSR
Number of runs	10

Table 3 Parameters of the proposed EOA-FCM technique

Simulation Parameters	Parametric values
Population size	40
k_m	0.5
β	-1.6
Convergence criterion	10^{-5}
No. of trial runs	30
Fuzzy membership	Trapezoidal
Fuzzy cluster design	Mandeni design

Simulation results of energy consumption are illustrated in figure 10 with the specified number of sensor nodes using three data aggregation methods. As shown in the graph, the developed new EOA-FCM approach consumes minimal energy than that of the conventional methods. Since, the developed EOA-FCM approach calculates the energy of each sensor node for data aggregation; the node residual energy is calculated to identify higher energy nodes. Based on the energy level, the nodes are grouped. Cluster head is selected for data gathering and is send to the base station. Due to this process, energy consumption while performing the data aggregation gets minimized significantly. The results of energy consumption with three different methods are compared and shown in figure 10. It is inferred from figure 8 that the EOA-FCM method minimizes the amount of energy for data aggregation by 9% and 19% as compared to existing MAEF approach [1] and SRS-CCDG approach [2].

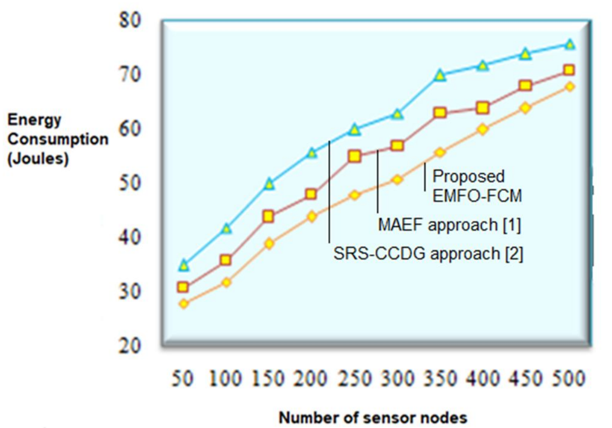


Figure 8 Comparative plots for energy consumption

TABLE IV. EVALUATED NETWORK LIFETIME

Number of sensor nodes	Network lifetime (%)		
	MAEF technique [1]	SRS-CCDG technique [2]	Proposed new EOA-FCM technique
50	82	78	88
100	86	83	90
150	85	81	89
200	86	83	91
250	88	85	90
300	84	80	89
350	87	84	91
400	83	80	89
450	89	85	93
500	87	83	91

I. Network lifetime

Network lifetime is measured as the ratio of number of energy efficient sensor nodes that are selected for data aggregation to that of the total number of sensor nodes. The network lifetime is calculated as follows,

$$N_{LT} = \left[\frac{Number\ of\ energy\ efficient\ sensor\ nodes\ are\ selected}{n}\right] * 100 \quad (8)$$

In equation (8),N_{LT} denotes the network lifetime, n represents the number of sensor nodes and the network lifetime is evaluated in percentage (%). Result of network lifetime is calculated employing the proposed EOA-FCM technique and that of other existing approaches MAEF [1] and SRS-CCDG [2] with the numbers of sensor nodes in the range of 50 to 500.In this case, 50 sensor nodes are considered for calculating the network lifetime. Among the 50 nodes, 44 nodes are selected to perform data aggregation. Network lifetime of developed new EOA-FCM technique is 88% whereas network lifetime of MAEF [1] and SRS-CCDG [2] are 82% and 78% respectively.

The network lifetime evaluated employing the developed EOA-FCM method is presented in table 4 based on the considered number of sensor nodes. Result confirms that the network life time increases significantly employing the new method than that of the conventional methods. The increase in network lifetime is due to the fact of identifying the optimal cluster unit using the proposed enhanced optimization algorithm and attaining the best cluster through which an effective data aggregation is carried out. Cluster head is selected for each cluster so as to maintain the energy efficient nodes. To protect the sensor nodes an effective data aggregation is of primary importance. It is also observed that the sensor nodes with minimal energy are unable to execute more tasks for a longer duration of time. Thus, the data aggregation with the new EOA-FCM technique has been an effective method to conserve and protect the resources in the cluster. This results in better network lifetime. Based on the results, it is well clear that the network life time is enhanced by 5% and 10% using the proposed method in comparison with that of the existing techniques from the literature [1, 2].

J. Data aggregation accuracy

Data aggregation accuracy is a metric which is a ratio of the number of data packets collected by sink node to that of total number of collected data points sent. This metric is evaluated using,

$$DAA\% = \left[\frac{Number\ of\ DP\ collcted\ by\ S_n}{Number\ of\ DP\ sent}\right] * 100 \quad (9)$$

In equation (9),DPdenotes the data packets, S_n represents the sink node and the accuracy metric is measured in percentage (%).

Simulation results evaluated for the metric data aggregation accuracy is shown in Figure 9 considering the number of data packets in the specified range of 25 to 250. Figure 9 presents the simulation results of the considered metric using the proposed technique and that of other compared methods. Among the compared three different data aggregation approaches, the developed new EAO-FCM technique hasreasonably attained better data aggregation accuracy (%). To determine the optimal number of clusters, the new EOA has been employed and the Bray-Curtis Similarity Index is applied to locate the similarity of cluster centroid and residual energy of sensor nodes. The resulting coefficient achieves higher similarity, and sensor node gets grouped into that cluster. The sensor node with higher residual energy is selected as cluster head. Cluster head collects the sensed data from sensor nodes within the cluster and will be sent to sink node with minimal energy consumption. Thus, sink node gathers data from the cluster head with minimal loss of energy. Due to this mechanism, higher data aggregation accuracy is achieved using the developed EOA-FCM technique and is increased by 5% and 11% in comparison with that of the previous techniques from existing literatures [1, 2].

K. Data aggregation time

Data aggregation time refers to amount of time consumed by sink node to gather data packets from the cluster head. The data aggregation time (T_{DA}) is computed using,

$$T_{DA} = n * t\ (aggregating\ one\ DP) \quad (10)$$

From equation (10), T_{DA}represents the data aggregation time, $n$specifies the number of data packets sent, t denotes the time for collecting one data packet,dp denote the data packets. The data aggregation time is measured in milliseconds (ms).

Table 5 presents the simulated results attained for data aggregation time with respect to the number of data packets employing the proposed EOA-FCM technique. It is inferred from Table 5, that the proposed EOA-FCM technique improves the energy efficiency of the data aggregation with minimum time in comparison with that of the MAEF [1] and SRS-CCDG [2] approaches respectively. The traditional data aggregation techniques collect the data from each sensor node and this leads to delay in data collection of the network. Employing the proposed EOA-FCM technique, the sink node collects the data from cluster head instead of collecting all the nodes within the network. This mechanism minimizes the data aggregation time due to the identification of best clusters using enhanced optimization technique and as well the applicability of Bray-Curtis similarity index for attaining the suitable similarity relation.

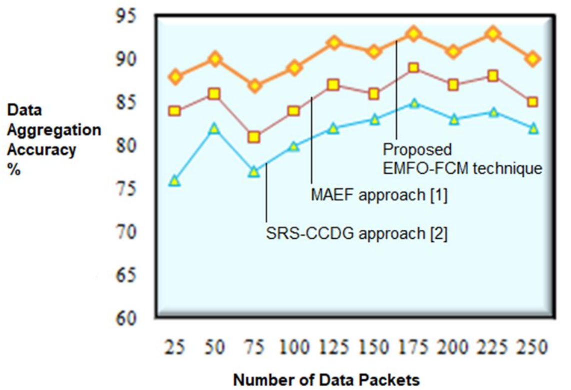


Figure 9 Comparative plots for data aggregation accuracy

TABLE V. EVALUATED DATA AGGREGATION TIME

Number of data packets	Data aggregation time (ms)		
	MAEF technique [1]	SRS-CCDG technique [2]	Proposed new EOA-FCM technique
25	20	25	15
50	23	26	19
75	27	32	23
100	28	35	25
125	34	38	29
150	39	45	33
175	44	47	39
200	52	56	46
225	56	61	50
250	58	63	53

During the simulation process, on consideration of 25 data packets the data aggregation time using EOA-FCM technique is 15ms whereas the data gathering time of MAEF[1] and SRS-CCDG [2]are 20ms and 25ms respectively. This elucidates that the results obtained using proposed EOA-FCM technique reduces the data aggregation time by 14% and 24% than the compared traditional techniques [1, 2]. It is well clear from the above comparative discussions, that the proposed Enhanced EOA-FCM approach has significantly minimized the energy consumption and enhanced the network lifetime proving its superiority over the other existing traditional techniques considered for comparison.

L. Statistical analysis and validation

In this paper, a novel enhanced moth-flame optimization based fuzzy c-means clustering algorithm is developed and simulated for the considered sensor network model. The classic moth-flame optimization considers only one way movement of the flame position whereas in the proposed enhanced moth-flame optimization technique, two best flame positions are taken and their mean values are evaluated and included in the distance evaluation mechanism. Additionally, in proposed EOA technique the worst moth position is included so as to improve the solution diversity. This feature of excluding worst moth position and including mean flame position improves the exploration and exploitation mechanism in large and tends to attain better optimal solution of the clusters.

The best optimal clusters is identified using the new EOA technique and then the elbow based fuzzy c-means clustering mechanism is employed to transmit and send the data packets between the cluster nodes and then towards the sink nodes. This process enhances the data aggregation mechanism and the applicability of Bray-Curtis similarity index tends to attain better similarity relation thereby minimizing the energy consumption and improving the lifetime of the network. The proposed EOA -FCM technique is stochastic population oriented approach and hence there exist random generation of moth population and membership definitions. Due to which, the proposed technique has to be statistically validated to prove its efficacy. This is done by determining the correlation coefficient (φ) and determination parameter (χ). When both these values are closer to 1, this substantiates the validity of the proposed model. On finding these values for the developed new EOA - FCM approach, the values evaluated are as given in Table 6.

TABLE VI. STATISTICAL ANALYSIS

Proposed technique	Parameter correlation coefficient (φ)	Parameter determination coefficient (χ)
New EOA-FCM approach	0.9927	0.9996

TABLE VII. MEAN SQUARE ERROR

No. of iterative runs	Mean Square Error
10	2.967521
20	2.341935
30	1.765449
40	1.554910
50	1.321872
60	0.993563
70	0.065213
80	0.031890
90	0.002187
100	0.0008321
120	0.0001259
124	0.0000764

It is clear from Table 6, that the values of $\varphi$and χ are nearer to 1, proving the validity of the proposed technique.

The mean square error as evaluated using equation (6) is given in Table 7 proving the reduction of this fitness function over the iteration process. Thus, the newly modeled EOA-FCM technique has proven its efficacy and superiority over the other methods compared for comparison. For attaining the best solution, this proposed technique has elapsed 124 iterations with a minimal mean square error of 0.0000764. This was taken to be the convergent point for attainment of solution using the proposed approach. Table 7 provides the mean squared error evaluated during the progress of the proposed new EOA -FCM approach.

IV. Conclusion

A novel enhanced moth-flame optimization based fuzzy c-means clustering algorithm has been designed and developed in this paper to perform energy efficient data aggregation of the considered sensor network models. The proposed model combining the advantages of the new EOA and that of elbow based FCM technique has achieved better data aggregation of the sensor network model. The metrics evaluated energy consumption, network lifetime, data aggregation accuracy and data aggregation time has attained better results using proposed technique than that of the other methods compared from literature. The superiority of the proposed technique is based on its attainment of best optimal cluster, selection of cluster heads and thereby increasing the lifetime of the network model. The hybridization of enhanced moth-flame optimization with elbow based fuzzy c-means clustering has resulted in better energy efficient data aggregation due to the combined features of both the methods. Also, employing this technique the data sensed from the cluster head gets aggregated by the sink time within a short span of time. Simulation results attained and its comparisons made had proven the superiority and effectiveness of the proposed EOA -FCM technique with that of the earlier other methods compared from literatures.

Conflict of Interest

The authors declare that there exists no conflict of interest in publishing this work.

References

[1] Mohamed El Fissaoui, AbderrahimBeni-hssane and MostafaSaadi, “Multi-mobile agent itinerary planning-based energy and fault aware data aggregation in wireless sensor networks”, EURASIP Journal on Wireless Communications and Networking, Springer, Volume 2018, Issue 92, Pages 1-11,2018

[2] Peng Sun , Liantao Wu , Zhibo Wang , Ming Xiao , Zhi Wang, “Sparsest Random Sampling for Cluster-Based Compressive Data Gathering in Wireless Sensor Networks”, IEEE Access , Volume 6, 2018, Pages 36383 – 36394

[3] ShimaPakdamanTirani, Avid Avokh, “On the performance of sink placement in WSNs considering energy-balanced compressive sensing-based data aggregation”, Journal of Network and Computer Applications, Elsevier, Volume 107, 2018, Pages 38-55

[4] CuicuiLv, Qiang Wang, Wenjie Yan, Jia Li, “A sparsity feedback-based data gathering algorithm for Wireless Sensor Networks”, Computer Networks, Elsevier, Volume 141, 4 August 2018, Pages 145-156

[5] Fen Zhou , Zhenzhong Chen , Song Guo , Jie Li, “Maximizing Lifetime of Data-Gathering Trees With Different Aggregation Modes in WSNs”, IEEE Sensors Journal , Volume 16, Issue 22, 2016, Pages 8167 – 8177

[6] DivyaLohani and ShirshuVarma, “Energy Efficient Data Aggregation in Mobile Agent-Based Wireless Sensor Network”, Wireless Personal Communications, Springer, Volume 89, Issue 4, 2016, Pages 1165–117

[7] Shusen Yang, UsmanAdeel, YadTahir, and Julie A. McCann, “Practical Opportunistic Data Collection in Wireless Sensor Networks with Mobile Sinks”, IEEE Transactions on Mobile Computing, Volume 16, Issue 5, May 2017.

[8] Soo-HoonMoona, Sunju Park, Seung-jae Han, “Energy efficient data collection in sink-centric wireless sensor networks: A cluster-ring approach”, Computer Communications, Elsevier, Volume 101, 2017, Pages 12-25

[9] Yongmin Zhang, Shibo He and Jiming Chen, “Near Optimal Data Gathering in Rechargeable Sensor Networks with a Mobile Sink”, IEEE Transactions on Mobile Computing, Volume 16, Issue 6, June 2017, Pages 1718-1729

[10] Diwakaran, B. Perumal, K. Vimala Devi, “A cluster prediction model-based data collection for energy efficient wireless sensor network”, The Journal of Supercomputing, Springer, Volume 75, Issue 6, 2019, Pages 3302–3316

[11] Ying Zhou , Lihua Yang , Longxiang Yang, and Meng Ni, “Novel Energy-Efficient Data Gathering Scheme Exploiting Spatial-Temporal Correlation for Wireless Sensor Networks”, Wireless Communications and Mobile Computing, Hindawi, Volume 2019, May 2019, Pages 1-10

[12] Rajesh K. Yadav , Daya Gupta, D. K. Lobiyal, “Energy Efficient Probabilistic Clustering Technique for Data Aggregation in Wireless Sensor Network”, Wireless Personal Communications, Springer, Volume 96, Issue 3.

[13] FereshtehKhorasani and Hamid Reza Naji, “Energy-efficient data aggregation in wireless sensor networks using neural networks”, International Journal of Sensor Networks, Volume 24, Issue 1, 2017, Pages 26-42

[14] SantoshSoni and Manish Shrivastava, “Novel Learning Algorithms for Efficient MobileSink Data Collection Using Reinforcement Learning in Wireless Sensor Network”, Wireless Communications and Mobile Computing, Hindawi, Volume 2018, August 2018, Pages 1-13

[15] Huthiafa Q. Qadori , Zuriati A. Zulkarnain , ZurinaMohdHanapi and ShamalaSubramaniam, “A Spawn Mobile Agent Itinerary Planning Approach for Energy-Efficient Data Gathering in Wireless Sensor Networks”, Sensors, Volume 17, Issue 6, 2017, Pages 1-16

[16] Zhetao Li, Yu Xin Liu, Ming Ma, Anfeng Liu, Xiaozhi Zhang, GungmingLuo, “MSDG: A novel green data gathering scheme for wireless sensor networks”, Computer Networks, Elsevier, Volume 142, 2018, Pages 223-239

[17] Yong Lu and Na Sun, “A resilient data aggregation method based on spatiotemporal correlation for wireless sensor networks”, EURASIP Journal on Wireless Communications and Networking, Springer, Volume 157, 2018.

[18] Khalid A.Darabkh, SajaM.Odetallah, Zouhair Al-qudah, Ala’ F.Khalifeh , Mohammad M.Shurman, Energy-Aware and Density-Based Clustering and Relaying Protocol (EA-DB-CRP) for gathering data in wireless sensor networks”, Applied Soft Computing, Elsevier, Volume 80, 2019, Pages 154-166

[19] Mohamed El Fissaoui, AbderrahimBeni-Hssane and MostafaSaadi, “Mobile Agent Protocol based energy-aware data Aggregation for wireless sensor networks”, Procedia Computer Science, Elsevier, Volume 113, 2017, Pages 25-32

[20] Chao Sha, Jian-meiQiu, Shu-yan Li, Meng-ye Qiang and Ru-chuan Wang, “A type of energy-efficient data gathering method based on single sinks moving along fixed points”, Peer-to-Peer Networking and Applications, Springer, Volume 11, Issue 3, May 2018, Pages 361–379

[21] Mirjalili, S., 2015. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. Knowledge-based systems, 89, pp.228-249.

[22] Yıldız, B.S. and Yıldız, A.R., 2017. Moth-flame optimization algorithm to determine optimal machining parameters in manufacturing processes. Materials Testing, 59(5), pp.425-429.

[23] Ho, Y.C. and Pepyne, D.L., 2002. Simple explanation of the no-free-lunch theorem and its implications. Journal of optimization theory and applications, 115(3), pp.549-570.