

Clustering and data aggregation scheme in underwater wireless acoustic sensor network

Vani Krishnaswamy*, Sunil Kumar S. Manvi

School of Computing and Information Technology, REVA University,
Bangalore-560064, India, tel: 080-66226622/ 080-28478539

*Corresponding author, e-mail: vanikrishnaswamy@gmail.com

Abstract

Underwater Wireless Acoustic Sensor Networks (UWASNs) are creating attentiveness in researchers due to its wide area of applications. To extract the data from underwater and transmit to watersurface, numerous clustering and data aggregation schemes are employed. The main objectives of clustering and data aggregation schemes are to decrease the consumption of energy and prolong the lifetime of the network. In this paper, we focus on initial clustering of sensor nodes based on their geographical locations using fuzzy logic. The probability of degree of belongingness of a sensor node to its cluster, along with number of clusters is analysed and discussed. Based on the energy and distance the cluster head nodes are determined. Finally using similarity function data aggregation is analysed and discussed. The proposed scheme is simulated in MATLAB and compared with LEACH algorithm. The simulation results indicate that the proposed scheme performs better in maximizing network lifetime and minimizing energy consumption.

Keywords: belongingness, clustering, euclidean distance, SSE, UWASN

Copyright © 2019 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

The rapid growth of research in underwater environment is due to numerous underwater communication applications such as oceanographic data collection, disaster prevention, undersea exploration, surveillance applications [1-3]. The unique challenges faced by UWASNs include large propagation delay (1.5x10 m/s), low bandwidth (KHz), high bit error rates, high mobility and difficulty in recharging the battery when compared to wireless communication in the terrestrial network.

The key issues in network topology are to improve the network lifetime and sustainability. Clustering and data aggregation schemes assist to accomplish this by making the network smaller and stable [4, 5]. The fundamental notch behind clustering is to divide the network into smaller units and logically organize the units to manage them easily. Clustering helps in reducing the communication overhead, energy efficiency, on the whole consumption of power, and increasing the lifetime of the network [6]. Researchers are vigorously working on various network clustering issues such as several ways of clustering [7-9] optimizing the number of clusters, selection of cluster head [10, 11], communication among clusters [12], and data aggregation in clusters [13-15].

Data aggregation is defined as the process that accumulates the data to minimize the transmission of redundant data and transmit the aggregated data to the sink or the Base Station (BS). The main aim of the data aggregation process is to congregate the data from the sensors and transmit it to the BS with least latency [13-15]. Hence it minimizes the energy and increase the lifetime of the network.

We propose a scheme to perform cluster based data aggregation considering the parameters like energy and distance with the following steps:

- (1) Initial clustering is performed by using fuzzy logic. At the same time, number of clusters in the network are determined using Sum of Squared Error (SSE) parameter.
- (2) Based on the distance and energy level the Cluster Head nodes are selected.
- (3) Cluster head nodes act as aggregators.
- (4) Using the concept of similarity function with Euclidean distance, these aggregators transmit the aggregated data to Base Station (BS).

Our contributions in this paper in comparison to the existing works include the following:

- Development of a mathematical model to evaluate the probability of sensor node belongingness to form the initial clusters. The sensor nodes are deployed stochastically and the nodes positions are stationary according to their communication range.
- Designing fuzzy clustering scheme based on the developed mathematical model.
- Determining the number of clusters using SSE parameter.
- Designing an energy level scheme for cluster head selection in the hierarchical topology.
- Using similarity function, design data aggregation scheme for transmitting the aggregated data to the BS.
- Comparative analysis of proposed scheme with LEACH algorithm [16] in terms of network lifespan and death rate of nodes.

The rest of the paper is organized as follows. Section 2 breifs about works related to various clustering schemes, selection of cluster head and several cluster based data aggregation schemes in networks. Section 3 explains the network model and energy model for cluster formation respectively. It explains the proposed fuzzy scheme in clustering, the cluster head selection considering parameters like energy level and distance and data aggregation scheme using similarity function. Section 4 deals with simulation and its parameters. Section 5 deals with result analysis. Summary of the proposed work and future works are presented in Section 6.

2. Related Works

In [17], the authors have proposed an agent based routing protocol which is energy efficient. The process of dynamic clustering is initiated and the cluster head along with the agents is responsible for data aggregation at the affected area. They have proposed an algorithm to increase the connectivity and reliability in the network. In [18], the authors have used Fuzzy Clustering Means (FCM) to select the cluster heads from an optimal number of clusters and setup the Underwater Isomorphic Sensor Network (UWISN). In addition to this, a scheme for determining the real cluster heads and selecting them have also been proposed. The authors in [19] have proposed a routing protocol based on grid with fuzzy logic where the entire network is separated into various virtual grids. Every grid in the network has only one active node which is selected using the fuzzy logic system.

In [20], the authors have proposed a new GPS-free routing protocol with Distributed Underwater Clustering Scheme (DUCS) which utilizes data aggregation to remove the redundant information and reduce the data loss in UWASN. In [21], the authors have proposed a scheme for an optimal selection of cluster head and cluster size using fuzzy logic along with inter and intra cluster communication considering the energy and the multiple paths for UWASN.

The authors in [22-24] have recommended clustering and aggregation techniques in UWASNs which are based on fuzzy logic system that captivates the residual energy, the node density, the link quality, the load and the distance to the sink/Base Station (BS) node. In [25], the authors have focused on designing an energy efficient routing protocol to transfer the data between sensor nodes utilizing the fixed courier nodes inorder to enhance the lifetime and decrease the end to end delay in the network. The authors in [26] have analyzed the consumption of energy in UWASNs for various transmission mechanisms with effect of changing ambient conditions.

3. Clustering

This section presents network model, clustering terminology and energy model employed for designing the proposed clustering scheme. It includes proposed fuzzy clustering scheme with selection of cluster head and data aggregation scheme.

3.1. Network Model

The network model explained here is similar to that presented in [16, 21] with the subsequent features. The sensor nodes are placed randomly in an underwater environment to form a 3-D static network where the communications between the sensor nodes are full duplex. The 3-D position information of each sensor node is achieved by positioning algorithms or by the use of hardware units, which are detected by acoustic waves.

The sensor nodes in the network are **homogeneous** and transmit the required information with different ranges of communication radius. The position of the sink node or BS is usually on the surface of the sea. The energy possessed by the BS is unlimited and it can communicate using underwater acoustic waves and radio waves. The processing and aggregation of data at the BS is carried out by each sensor node. During this process, the energy of the sensor node is reduced gradually; this in turn results in a dead node. The network is considered to be dead when the dead nodes number exceeds beyond the threshold limit. Therefore, the aim of the hierarchical topology scheme is to increase the network lifetime to the maximum feasible extent.

3.2. Energy Model

The major limitation of UWASN is its energy capability to recharge the battery which cannot be done frequently. In underwater environment the consumption of energy by the nodes depends on the following factors. 1) To sense, receive and process the data. 2) To convey the collective data to the sink. The first factor is considered, as the energy consumption is less when compared to the second. This model of energy consumption for transmitting data by the nodes is presented in [24]. The Unit energy consumed to process one bit of message is denoted as $E_{tx}(d)$ and is evaluated according to the (1):

$$E_{tx}(d) = P_r \times T_p \times A(d) \quad (1)$$

where P_r represents the threshold power of a node for receiving the data package, d represents the distance of transmitting data package, and T_p represents the time of the transmitting data package which is represented as follows:

$$T_p = \frac{M_b}{S_v}$$

where M_b and S_v represents the size and the transmission speed of data package respectively. The energy attenuation $A(d)$ with the transmitting distance of the data package is 'd' is calculated as follows:

$$A(d) = d^\lambda \times \beta d$$

where λ represents the energy spreading factor which is 1 for cylindrical, 1.5 for practical and 2 for spherical spreading respectively. The parameter

$$\beta = 10 \frac{\alpha(f)}{10}$$

absorption coefficient $\alpha(f)$, which can be calculated using the following equation:

$$\alpha(f) = 0.11 \frac{10^{-3}(f^2)}{1+f^2} + 44(10^{-3}) \frac{f^2}{(4100+f^2)} + 2.75 \times 10^{-7} f^2 + 3 \times 10^{-6}$$

where 'f' is the frequency of the carrier acoustic signal in KHz, and $\alpha(f)$ is in dB/m. Considering the underwater environment, the amount of energy consumed to transmit 'l' bit of data over a distance d by the node is given by $E_{tx}(d, l)$ as shown in the (2):

$$E_{tx}(d, l) = l E_{elec} + l \times T_p \times C \times H \times d \times \beta d \quad (2)$$

where H represents the depth of the node in *mtrs*.

$$C = (2\pi (0.67)109.5)$$

The amount of energy consumed to receive 'l' bit of data by the receiver is given by $E_{rx}(d, l)$. To be specific the threshold value 'd0' is set, which is related to the transfer distance. If the transfer distance is less than 'd0', then the energy consumption proportional to the

squared distance else, the energy consumed is proportional to the fourth power of the distance. The energy consumption for transmitting and receiving 1 bit of data with the distance 'd' is calculated in (3) and (4):

$$E_{tx}(d, 1) = \begin{cases} 1.E_{elec} + 1. \in fs. d^2 & \text{if } d < d_0 \\ 1.E_{elec} + 1. \in fs. d^4 & \text{if } d \geq d_0 \end{cases} \quad (3)$$

$$E_{rx}(d, 1) = 1 + E_{elec} \quad (4)$$

where E_{elec} represents the **electronics energy**, which depends on the energy dissipated per bit to run the transmitter or the receiver. The energy cost of the signal amplifier in two communication modes considering the distance between transmitter and receiver are represented as $\epsilon fs: d^2$, $\epsilon fs: d^4$. In this paper the assumption made regarding data fusion is that in spite of the number of nodes in a cluster, every node in a cluster gathers and transmits l bit to the cluster head, and in turn the cluster head condenses the overall received information to l bit.

3.3. Proposed Fuzzy Clustering Scheme

The cluster formation is performed by taking into account the given 3-D network environment. Cluster structure is characterized by two types of nodes called *Member cluster nodes* and *ClusterHeadnodes* which are considered as the backbone of the network. The *Memberclusternodes* that are connected to its own *ClusterHeadnode* lie dormant to save the energy consumption of the network. Whereas the *ClusterHeadnodes* that are connected to the closest neighbor node of other clusters is usually in a shallower location. The process of selecting the *ClusterHeadnodes* and the reconstruction of the network is called as a *round*. This process is achieved periodically to decrease the consumption of energy in the network which in turn increases the network's lifetime.

Before the communication begins between the sensor nodes, initially the nodes are divided according to locations of the **nodes into a number of fuzzy subsets** using **fuzzy clustering model**. During the formation of clusters, the nodes which are closer to the location and which require less energy for communication are assigned to the same cluster. Initially at the commencement of each round, which is based on assured probabilities, every node in the network belongs to the initial subsets of the clusters. The process of selection of the cluster heads will be performed by each subset in parallel and the messages are transmitted to the base station. The size of the network and time required for the calculation are reduced in this scheme.

A method is proposed to **segregate** the underwater acoustic wireless sensor nodes into **m primary fuzzy clusters**. Further, the optimal numbers of clusters are obtained using an **elbow method**. The next step, determines the matrix w_{ij} ; $1 \leq i < n$; $1 \leq j < m$ by means of fuzzy clustering method. Prior to every round, each node is related to a suitable cluster depending on the degree of belongingness of each node. If the node y_i is applied to B_j , then it satisfies the condition given in (5).

$$\sum_{l=1}^{j-1} W_{il1} \leq r < \sum_{l=2}^j W_{il2} \quad (5)$$

The random number with uniform distribution is between 0 and 1 is denoted by r . The basic definition of traditional clustering is the division of basic set of identical objects into numerous subsets. In reality, the real clusters which are formed are typically much complex and the dependency of objects to them is more fuzzy. In short, such clusters are called fuzzy subsets of the basic set which are comprised of delicate parts. This is how, the concept of fuzzy clustering was proposed. In this paper, the process of initialization is proposed considering the fuzzy clustering model. Basically in this clustering method, the set of n objects are partitioned $A = a_1, a_2, a_3 \dots a_n$ into m fuzzy clusters $B_1, B_2 \dots B_m$, the clustering is denoted as nm matrix:

$$C = [w_{ij}](1_i < n, 1_{jm})$$

where w_{ij} = the degree of belongingness of the i th object to the j th cluster. The matrix $C = [w_{ij}]$ must convince the subsequent conditions. For each object a_i and cluster B_j , $0 \leq w_{ij} \leq 1$

for each object a_i , $\sum_{j=1}^m w_{ij} = 1$. For each cluster B_j , $0 \leq \sum_{i=1}^n w_{ij} < n$. To calculate the center of the cluster B_j , $1 \leq j \leq m$ is b_j . The degree of belongingness of object of a_i to cluster b_j is expressed in terms of the distance between the a_i and the center of the cluster b_j which is represented as follows $\text{dist}(a_i, b_j)$. The possibility of any object belonging to the cluster can be determined by the distance between the object and the center of cluster. For example shorter the distance between the cluster center b_j and the object x_j , greater are the opportunities for the object x_j belonging to the corresponding cluster B_j . The amount of belongingness of an object a_i to cluster B_j is expressed using the (6).

$$w_{ij} = \frac{1}{\text{dist}(x_i, b_j)^2} \quad (6)$$

The definition of degree of belongingness w_{ij} is obtained by normalizing the (6) which satisfies the conditions of matrix $C = [w_{ij}]$.

$$w_{ij} = \frac{\frac{1}{\text{dist}(x_i, b_j)^2}}{\sum_{l=1}^m \frac{1}{\text{dist}(x_i, b_l)^2}} \quad (7)$$

Clustering algorithms are broadly classified into hard and soft clustering algorithms. Fuzzy clustering belongs to **soft clustering** because in this algorithm every object belongs to **multiple clusters**. The following are the advantages of soft clustering. 1) Every object belongs to multiple clusters; hence the user can observe multiple themes for a cluster. 2) Various clusters get formed for various themes. 3) In order to calculate the order of the object appropriately, the measure related between clusters and objects can be used as a relevance measure. The initialization of fuzzy clustering is based on expectation maximization algorithm [23]. This expectation maximization algorithm used for fuzzy clustering generates ' m ' different clusters are briefed in the steps which is given in Process 1:

- 1) Classify the set of ' n ' objects based on its features into ' m ' clusters.
- 2) The cluster centers B_j are selected by uniform distribution of random vectors.
- 3) Compute the degree of belongingness (w_{ij}) using the distance between the nodes within the cluster.
- 4) Normalize (w_{ij}) until we get ' m ' clusters

3.4. Determination of the Number of Clusters

To decide the suitable cluster number is a cumbersome task especially in fuzzy clustering. The granularity of the clustering and the discovery of an appropriate balance between precision and compressibility have to be managed. The sum of squared error (SSE) for each cluster is defined in (8).

$$SSE(C_j) = \sum_{i=1}^n w_{ij}^p \text{dist}(a_i, b_j)^2 \quad (8)$$

$$SSE(m) = \sum_{i=1}^n \sum_{j=1}^m w_{ij}^p \text{dist}(a_i, b_j)^2 \quad (9)$$

The SSE inside each cluster can be decreased by increasing the number of clusters. The above concept results in better characters of the data objects which are retained from a number of clusters, in a manner such that the objects in the cluster are more analogous to each other. There will be a trivial reduction in SSE in each cluster due to the splitting of cluster into sub clusters. To decide on the number of clusters, Elbow method which is an efficient algorithm is employed. When the number of clusters $m > 0$; the following steps are followed:

- 1) Determine the $SSE(m)$;
- 2) Sketch the curve between the determined $SSE(m)$ and the variable ' m ';
- 3) The accurate number of cluster will be implied from the most significant inflection point on the curve;

where p ($p \geq 0$) represents a parameter to determine the priority weighting of the degree of belongingness w_{ij} . The amount of fitness of the data can be evaluated using the SSE for fuzzy clustering with 'm' clusters as defined in (9).

3.5. Selection of Cluster Head

According to the analysis of energy consumption, there are few disadvantages of using LEACH protocol in underwater environment.

- The energy of the nodes which are distant from the sink/BS is exhausted early.
- The cluster head nodes are randomly selected and gets concentrated due to which the energy efficiency of the node will decrease.

To avoid the disadvantages of the LEACH protocol [16] and to select the cluster head node with high energy and to enhance the life time of network, a new scheme of selecting the cluster head node is proposed by taking the idea from the work given in [8]. The sink/BS will broadcast the information; the nodes are classified into different levels based on the strength of the information and the distance from the sink as level 1, level 2 as shown in the Figure 1. The node nearer to the base station has the lesser level number and more chance of selecting as CH (Cluster Head).

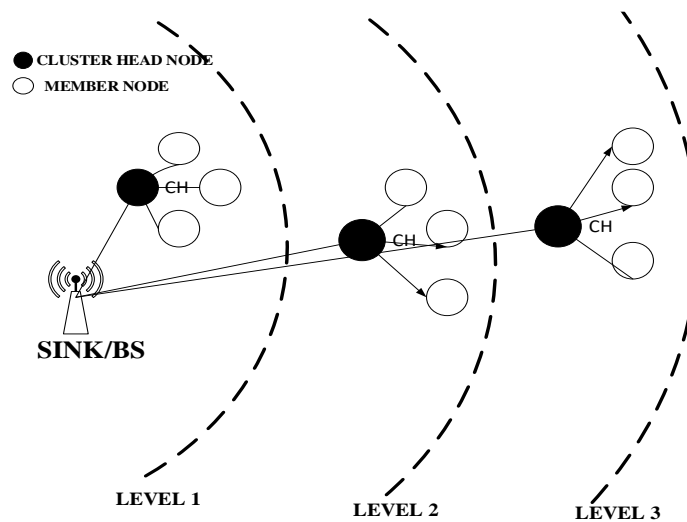


Figure 1. Energy level classification

The wait time of the node is represented using the formula:

$$W_T = \text{init} \left[N \left(1 - \frac{E_r}{E_i} \right) \right] + (L_i - 1) + I_D^i / N \quad (10)$$

where E_r = residual energy of node i ; E_i = initial energy of node i ; N = number of nodes; L_i = level number of node i ; I_D^i = Identification number of node i . From the (9), it is to infer that the member node with more energy will broadcast the message more quickly when compared to other nodes. If two or more nodes have same energy in different energy levels, then the node with higher energy level is chosen primarily to transmit the message. Process 2 explains about the steps involved in selecting the cluster head node. Process 2:

The following are the steps followed to select the Cluster Head (CH) node.

- 1) Initially during setting up of network, base station broadcasts the messages to all the nodes. Every node will know their energy Level number L_i using the strength of the received power. Then, calculate T_i using L_i and energy level.
- 2) Nodes (m for every round) with higher T_i are selected as cluster head nodes. Accordingly using CSMA (Carrier Sense Multiple Access) MAC (Medium Access Control) protocol CH will broadcast advertisement message (ADV) in time with T_i .

- 3) Depending on the strength of the received signal, every member node other than CH will determine its CH for the next round.
- 4) Once again using CSMA MAC protocol every non CH node will transmit a join-request back to its chosen CH.
- 5) Using TDMA (Time Division Multiple Access) CH node will schedule for data transmission within the cluster.

The uniform distribution of CH for the entire network is ensured, when a node receives the strong signal of ADV message, it will surrender the opportunity to turn out to be CH which avoids the CH to get close.

3.5. Data Aggregation Scheme

The cluster head periodically receives the information from the member nodes in the network. The collective data received by the CH will be transmitted to the sink. Sequentially to avoid the data redundancy which results in duplication of data and reduce the energy consumption in transmissions, a data aggregation scheme is proposed taking the idea from the work given in [14]. A data aggregation scheme has been implemented among CHs using the concept of similarity function. Euclidean distance formula is used with similarity function. CH collects all the data transmitted from its member nodes and accumulates as a set of data called *vector*. The comparisons of two vectors are performed using similarity function and if two vectors are found to be alike then CH will transmit only one data in place of both to the sink. This process avoids the data redundancy which in turn reduces the energy consumption in the network.

4. Simulation

This section presents simulation model, simulation parameter inputs and performance parameters.

4.1. Simulation Model

A node is considered to be dead or alive depending on the available energy. If the node energy reduces to 0, then it is considered as a dead node. Simultaneously in the network, if the count of dead nodes exceeds a cut off value, then the entire network is said to be deceased. Network environment discussed in Section 3 is simulated for analyzing the performance of clustering scheme.

The simulations were carried out using MATLAB and the performances of LEACH and the proposed FBC algorithm were analyzed in terms of the number of dead and alive nodes, Number of cluster using SSE and the total energy. The sensor nodes were randomly deployed in the region S and are able to communicate with each other. We assume that the base station is built at two different positions ($25\text{ m} \times 25\text{ m} \times 50\text{ m}$) and ($50\text{ m} \times 50\text{ m} \times 100\text{ m}$). For simulations, the number of sensor nodes N set were 100 and with data packet size of 400 bits for every transmission time. The initial energy of each node and electronic energy set are 0.5 J and 50 nJ/bit respectively. In addition to this, the energy of the base station is assumed to be unlimited as it is solar powered.

4.2. Simulation Parameters

The simulation inputs are shown in Table 1. The performance parameters are explained as follows.

- Number of clusters: Elbow method which is an efficient algorithm is employed to find out the number of cluster using SSE parameter.
- Life cycle: The study of life cycle of the nodes with respect to their initial energy and position of the base stations are conducted.
- Energy consumption: In data aggregation phase the CH aggregates the data and transmits to the sink. The comparative analysis of consumption of energy with and without data aggregation is performed.

Table 1. The Parameters for Simulation

Variable	Parameter	Value
S	Distribution area	100x100x100m ³
n	Number of nodes	100
l	length of every data	400 bit package
E _{elec}	Energy cost of data aggregation	10mJ/bit
E _{DA}	Factors in objective function	5mJ/bit
α, β, γ		0.2,0.3,0.5

5. Result Analysis

This section presents the comparative analysis of proposed scheme (denoted as FBC) and LEACH algorithm [16].

5.1. Determination of Cluster Amount

Experiments were carried out for 50 iterations using process (1). The value of SSE within each cluster was determined by the variable m . Figure 2 depicts the relationship between SSE and the number of clusters. The graph shows that when the number of cluster amounts to 2, SSE has higher value. Consequently, number of clusters set to 2 for conducting all the experiments.

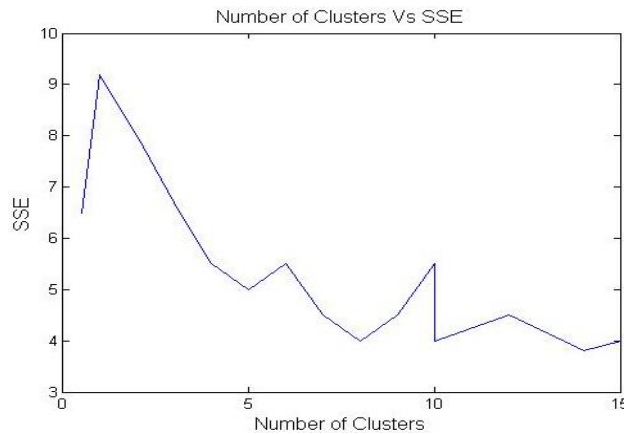


Figure 2. Relation between numbers of clusters vs SSE

5.2. Study of Life Cycle

The life cycle of the UWASN is greatly influenced by two factors. (1) The initial energy of the nodes. (2) The position of the Base Station (BS). In the experiment conducted it was decided to have two different situations: first considering the first factor where every node in the network has the same initial energy as 0.5J and secondly uniformly distributing the energy among the nodes between 0.3 to 0.6 J.

The positions of the BS was varied at (25, 25, 50) which was nearer to the network area and at (50, 50, 100) which was considerably far from the network area. The entire experiment was repeated for LEACH algorithm and FBC. It was found that the death of a few nodes in the network area did not have an immense impact on network lifetime, especially when the redundancy of the network coverage was more. After that we considered dead nodes with network lifetime. Figures 3 and 4 depicts the relationship between the percentage of dead nodes and the rounds.

Figures 3 and 4 shows that the positions of the BS are different, irrespective of the fact that the network nodes contained an equal initial energy. In both the situations the proposed algorithm FBC achieved better results when compared to the LEACH algorithm resulting in prolonging the expiry of the nodes. The most important reason for this is that, in our proposed algorithm the selection of cluster heads were achieved effectively. This in turn decreased the consumption of the energy for the communication between the nodes. Simultaneously the

nodes with high outstanding energy had the preference of being cluster heads, which balanced the consumption of the energy in each network node and thereby avoiding the early death of the nodes. Hence the lifespan of the network was extended. It was also observed from the graph that the BS nearer to the network area performed better when compared to the BS located at a larger distance in the network area. In both the situations the proposed algorithm performed well when compared to LEACH.

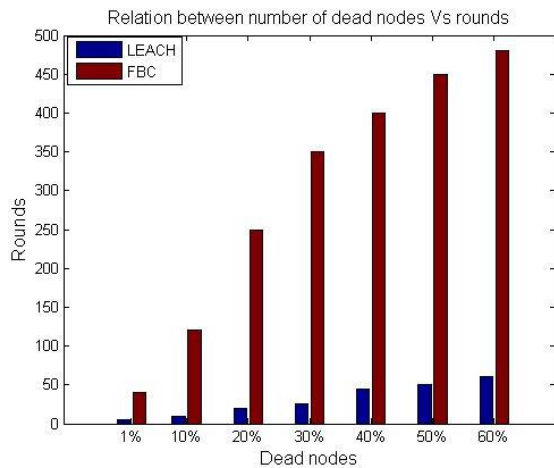


Figure 3. Dead nodes vs rounds with same energy at the BS (25, 25, 50)

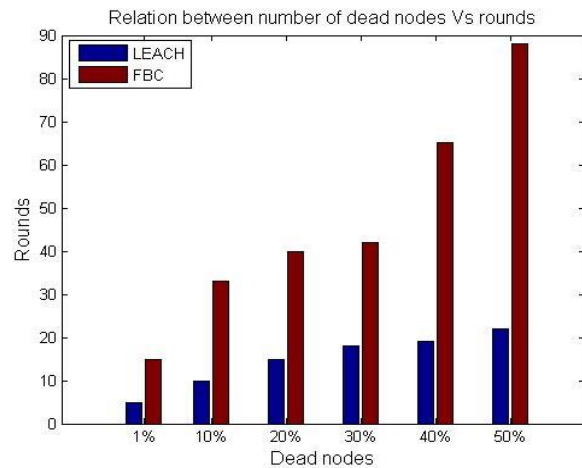


Figure 4. Dead nodes vs rounds with same energy at the BS (50, 50, 100)

Practically when we consider the optimal cost point, smaller range controls the redundancy of the network. During such situation, to assure that the entire network is connected, on a condition that all nodes in the network stay alive for a longer duration. Even when one node demise, the quality of service across the network is momentarily reduced resulting in focusing more interest on the living rate of UWASN.

Figures 5 and 6 depict the relation between rate of survivability of nodes and the network life cycle. The graphs show that when compared to LEACH algorithm, FBC has a smaller curved slope which indicates that the process of nodes dying was reasonably placid. This is because in FBC, both the distance and energy are considered to share the energy consumption between each node. Thus assuring that none of the nodes in the network diminished their energy, ultimately the lifespan of the node was extended.

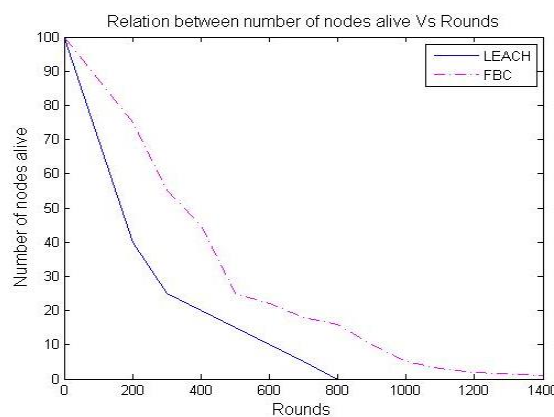


Figure 5. Nodes alive vs rounds with same energy at the BS (25, 25, 50)

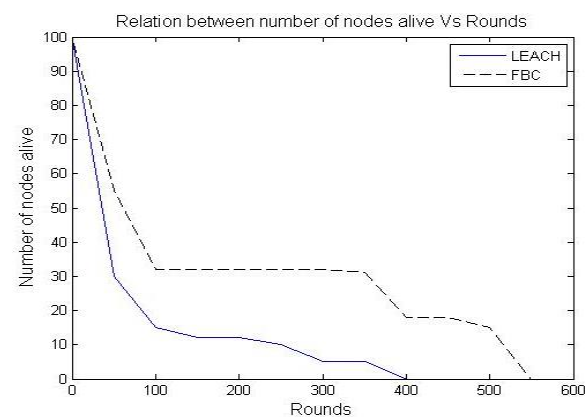


Figure 6. Nodes alive vs rounds with same energy at the BS (50, 50, 100)

Figure 7 depicts the consumption of the energy in the network, considering both the schemes of with and without data aggregation. The overall consumption of the energy in the network with and without data aggregation is indicated by red line and blue line respectively. In the proposed scheme of clustering and cluster head selection, the energy is saved at each phase. The data aggregation and transmission of aggregated data to the BS is performed by cluster head nodes. Also, the energy is saved in the data aggregation scheme as it uses similarity function through which it reduces the number of duplicate data transmissions from cluster-heads to the BS/sink. As a result, networks using clustering with data aggregation scheme devour less energy compared to a network without clustering and data aggregation.

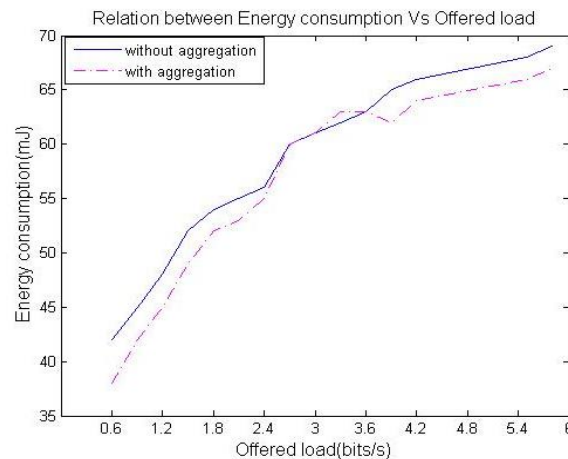


Figure 7. Energy consumption vs offered load

6. Conclusion

In this study, we propose a new clustering scheme using fuzzy logic considering energy and the probability of belongingness of the sensor nodes. The cluster head selection is performed based on the factors such as the energy and distance. Further, using similarity function the aggregated data at CH is transmitted to the BS. A simulation result shows that the proposed scheme performs better in prolonging the lifespan of the network. As a future enhancement, various PSO (Particle Swarm Optimization) alternatives could be used to resolve the problem of cluster head selection and analysis could be carried out based on their performances. In addition to this various initial clustering algorithms could be designed to decrease the redundant data. This would result in the energy being conserved and in turn increasing the lifespan of the network. We are focusing to work on various data aggregation schemes to achieve better accuracy of data.

References

- [1] Akyildiz TF, Pompili D, Melodia T. Underwater acoustic sensor networks: research challenges. *Ad Hoc Networks*. 2005; 3(3): 257-279.
- [2] PV Amoli. An Overview on Current Researches on Underwater Sensor Networks: Applications, Challenges and Future Trends. *International Journal of Electrical and Computer Engineering*. 2016; 6(3): 955.
- [3] IF Akyildiz, D Pompili, T Melodia. State-of-the-art in protocol research for underwater acoustic sensor networks. *WUWNet 06*. 2006: 7-16.
- [4] Rakesh Kumar and Navdeep Singh. A Survey on Data Aggregation and Clustering Schemes in Underwater Sensor Networks. *International Journal of Grid and Distributed Computing*. 2014; 7: 29-52.
- [5] Kumar R, Singh N, A survey on data aggregation and clustering schemes in underwater sensor networks. *Int. J. Grid Distrib. Comput.* 2014; 7(6): 29-52.
- [6] Yu JY, Chong PHJ. A survey of clustering schemes for mobile adhoc networks. *IEEE Communications Surveys and Tutorials*. 2005; 7(1): 32-48.

- [7] Domingo MC, Prior C. A distributed clustering scheme for underwater wireless sensor networks. *PIMRC* 07. 2007: 1-5.
- [8] Manjula RB, Manvi SS. Cluster based data aggregation in underwater acoustic sensor networks. *INDICON12*. 2012: 104-109.
- [9] Wang P Li, C Zheng J, Mouftah HT. A dependable clustering protocol for survivable. *ICC* 08. 2008: 3263-3268.
- [10] Yang G, Xiao M, Cheng E, Zhang J. A cluster-head selection scheme for underwater acoustic sensor networks. *CMC* 10. 2010: 188-191.
- [11] Buddha S, Lobiyal DK. A novel energy-aware cluster head selection based on particle swarm optimization for wireless sensor networks. *Humancentric Computing and Information Sciences* 12. 2012: 1-18.
- [12] Goyal N, Dave M, Verma AK. Energy Efficient Architecture for Intra and Inter Cluster Communication for Underwater Wireless Sensor Networks. *Wireless Personal Communications*. 2016; 89: 687-707.
- [13] Goyal N. Data aggregation in underwater wireless sensor network: Recent approaches and issues. *Journal of King Saud University-Computer and Information Sciences*. 2017; 4(7): 1-12.
- [14] Tran KTM, Oh SH, Byun JH. Well-suited similarity functions for data aggregation in cluster-based underwater wireless sensor networks. *International Journal of Distributed Sensor Networks*. 2013; 9(8): 1-7.
- [15] Virmani D, Sharma T, Sharma R. Adaptive energy aware data aggregation tree for wireless sensor networks. *International Journal of Hybrid Information Technology*. 2013; 6(1): 25-36.
- [16] WB Heinzelman, AP Chandrakasan, H Balakrishnan. Application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*. 2002; 1(4): 660-670.
- [17] Bharamagoudra MR, Manvi SS, Gonen B. Event driven energy depth and channel aware routing for underwater acoustic sensor networks: agent oriented clustering based approach. *Elsevier Journal of Computer and Electrical Engineering*. 2017; 58: 1–19.
- [18] Wang F, Wang L, Han Y, Liu B, Wang J, Su X. A Study on the Clustering Technology of Underwater Isomorphic Sensor Networks Based on Energy Balance. *Sensors*. 2014: 12523-12532.
- [19] Reza MM, Rahman KT, Zakaria Rakesh Kumar, Navdeep Singh. A.S.M. Grid based Fuzzy Optimized Routing Protocol for Underwater Sensor Networks. *Int. J. Comput. Appl.* 2015; 112: 2326.
- [20] Domingo MC, Prior R. *Design and Analysis of a GPS-free Routing Protocol for Underwater Wireless Sensor Networks in Deep Water*. 2007 International Conference on Sensor Technologies and Applications (SENSORCOMM 2007). 2007: 215-220.
- [21] Ni Q, Pan Q, Du H, Cao C, Zhai Y. A Novel Cluster Head Selection Algorithm Based on Fuzzy Clustering and Particle Swarm Optimization. *IEEE/ACM Transactions on computational Biology and Bioinformatics*. 2017; 14(1): 76-84.
- [22] Rakesh Kumar and Navdeep Singh. *Fuzzy Based Clustering and Aggregation Technique for Under Water Wireless Sensor Networks*. ICECS. 2014.
- [23] Nasser S, Alkhaldi R, Vert. *A modified fuzzy K-means clustering using expectation maximization*. Proc. IEEE Int. Conf. Fuzzy Syst. 2006: 231-235.
- [24] Quost B, Denux. *Clustering Fuzzy Data Using the Fuzzy EM Algorithm*. SUM '10. 2010: 333-346.
- [25] Ahmed Mukhtiar, Salleh Mazleen, Channa Muhammad, Rohani Mohd. RMEER: Reliable Multipath Energy Efficient routing protocol for underwater wireless sensor network. *International Journal of Electrical and Computer Engineering*. 2018; 8(6): 4366-4373.
- [26] Sehgal Anuj, David Catalin, Schnwlder Jrgen. *Energy Consumption Analysis of Underwater Acoustic Sensor Networks*. OCEANS'11-MTS/IEEE. 2011: 1-6.