Journal of Engineering Science and Technology
Vol. 14, No. 3 (2019) 1200 - 1219

ENERGY-AWARE ON-DEMAND FUZZY-UNEQUAL CLUSTERING PROTOCOL FOR WIRELESS SENSOR NETWORKS

D. R. DAS ADHIKARY*, DHEERESH K. MALLICK

Department of Computer Science and Engineering, Birla Institute of Technology,
Mesra, Ranchi-835215, Jharkhand, India
*Corresponding Author: dibya@bitmesra.ac.in

Abstract

Clustering is a mechanism by which, the network is partitioned into disjoint sets of groups, in order to achieve energy efficiency and facilitate data aggregation in a wireless sensor network. The clustering algorithm proposed earlier performs clustering on a round by round basis. On-demand clustering is a recent approach, which eradicates every round based clustering by performing clustering when it is required. However, the on-demand clustering approaches proposed so far use equality in clustering where clusters are of almost equal size thus, it suffers from the hot spot problem and unbalanced energy consumption problem. Therefore, in order to solve the hot spot problem and to create a balanced energy consumption structure, an energy-aware on-demand fuzzy-logic based unequal clustering approach is proposed. A comparative analysis of the proposed approach with Energy-aware distributed dynamic Clustering Protocol using Fuzzy-logic (ECPF) and Low-Energy Adaptive Clustering Hierarchy (LEACH) shows that the proposed approach performs far better, in terms of energy consumption and lifetime metrics.

Keywords: Clustering, Fuzzy-logic, On-demand clustering, Unequal clustering, Wireless sensor networks.

1. Introduction

A group of sensors with wireless communication facility when deployed to observe an area of interest and facilitate report to an observatory system forms a network like structure, broadly known as Wireless Sensor Network (WSN). Typically, a sensor is a tiny device with built-in sensing and wireless communication facility [1]. The role of a sensor is confined to sense its vicinity, process the sensed data and send the data to the observatory system commonly known as the Base Station or sink [2].

The tininess of the sensor brings several limitations to its functionality. One such limitation is the battery. Sensor nodes typically operate over limited battery power and due to the hard-to-reach area of deployment, recharge or replacement of the battery is non-viable [3, 4]. However, the surveillance nature of applications required long-term functioning of the network. Thus, energy efficiency is a major concern in WSNs and Clustering approach is instrumental in achieving energy efficiency [5]. It is also helpful in data aggregation and network management [6].

In clustering, the whole network is divided into a disjoint set of nodes called clusters with a designated leader called Cluster Head (CH). The CH works as a local sink to collect data from its cluster member and send it to the sink. The operation of a typical clustering approach carried out in rounds. Most of the clustering algorithm proposed earlier select CH in every round. The action is essential in selecting the CHs with best resources [7]. However, the problem is, a clustering process requires many control information exchanges and every round based cluster formation process waste a large sum of energy. To solve this problem, few authors proposed on-demand based clustering. In on-demand, based clustering the clustering process is triggered only when it is essential for the CHs to give up its responsibility. However, the problem arises when the load of the CHs is unbalanced. Unbalanced load of CHs means unequal energy consumption rate. Therefore, the CHs, which have more energy consumption per round trigger the clustering process, thus diminish the benefit of the on-demand based clustering process. Unequal clustering is known for its load distribution properties [8]. In unequal clustering clusters close to the sink are smaller so that it can spend lesser energy on intra-cluster traffic and the preserved energy can be spent on inter-cluster relay traffic [9].

Fuzzy Inference System (FIS) has been successfully used in several scenarios in WSN [10, 11]. The usefulness of FIS in WSN is precisely due to two main reasons. Firstly, WSNs are resource scant network and FIS imposes less computational complexity on the system. Secondly, FIS can make the decision even if there is insufficient or incomplete information about the system and for a dynamic environment like WSNs, one cannot guarantee precise, unambiguous information all the time.

Therefore, in this paper, an on-demand clustering with unequal cluster range for CHs is proposed. The sole objective of the proposed approach is of maximising the life expectancy of WSNs. To achieve this goal, the proposed approach focused on selecting the best possible CHs in terms of energy and assigning load proportional to CH's remaining energy, distance to sink and centrality to its neighbour. Moreover, the proposed approach minimizes the wastage of energy in unnecessary clustering by sporadically selecting the CHs.

The remaining of the paper is summarised as follows: A brief summary of some of the well-known clustering approaches is discussed, followed by the system preliminaries. The proposed methodology is then discussed thoroughly. Subsequently, the detailed simulation and analysis of the proposed work along with all the results is discussed and finally, the paper is concluded.

2. Review of Literature

Over the last decade, many clustering algorithms have been proposed. LEACH [12] is the first protocol to address the energy issue in WSN by applying clustering. Although LEACH is able to extend the network lifetime, there are few problems exist with it, which are addressed by many successive approaches. Randomized selection of CHs in LEACH leads to unbalanced energy consumption structure. Therefore, in LEACH-C [13], the CHs are deterministically selected by the sink to ensure balanced clustering structure. Similarly, Xiangning and Yulin [14] proposed Multihop-LEACH and Energy-LEACH to solve the direct communication problem and randomized selection problem of LEACH. Hybrid Energy-Efficient Distributed (HEED) [15, 16] is another such instance where the residual energy of nodes is used for selecting the CHs.

In order to eradicate the unbalanced energy consumption problem, few authors proposed unequal clustering [8, 9, 17] as an effective way of balancing the energy consumption. Soro and Heinzelman [17] first discussed the principle of un-equality in clustering. They proposed a scheme to form adaptive clusters based on their distance to the sink. Li et al. [8] proposed another unequal clustering mechanism, which selecs CHs based on competition. Unequal Cluster-based Routing (UCR) [9] is another such approach, which adopts unequal clustering to balance energy consumption.

Some clustering algorithms utilize fuzzy-logic to resolve uncertainties in WSNs. First, Gupta et al. [10] used fuzzy-logic to deal with uncertainties in the selection of CHs. Subsequently, Kim et al. [18] and Anno et al. [19] make use of fuzzy logic in the selection of CHs. Bagci and Yazici [11, 20] and Sert et al. [21] proposed an unequal clustering approach, which uses fuzzy logic to assign distinct clustering range for CHs. Adhikary and Mallick [22] used fuzzy logic in a multi-hop communication scenario to create a stable route for relaying data.

On-demand clustering is a recent trend in clustering WSN where the cluster is formed when required. Taheri et al [7] proposed ECPF, which introduced on-demand, based clustering and it uses fuzzy-logic to select the CHs. Although ECPF minimizes the energy consumption, however, it is unable to proper balancing the network load.

Better energy efficiency can only be achieved through minimizing and balancing the energy consumption of the network. Therefore, in the proposed approach on-demand clustering similar to ECPF is used to minimize the unnecessary re-clustering of the network. However, for efficient load distribution, the proposed approach focused on energy-rich CH selection, assigning load proportional to the node's competence and distance based multi-hop inter-cluster communication. However, unlike Li et al. [8], Chen et al. [9], Gupta et al. [10], Bagci and Yazici [11, 20], Heinzelman et al. [12], Xiangning and Yulin [14], Soro and Heinzelmann [17], and Sert et al. [21] proposed

approaches, which use an energy based distributed approach in CH selection, however, unlike Li et al. [8], Chen et al. [9], Bagci and Yazici [11, 20], Heizelman et al. [13], and Sert et al. [21] proposed approaches, which use remaining energy, distance to sink and centrality to efficiently distribute network load and unlike Adhikary and Mallick [22], the multi-hop inter-cluster communication of the proposed approach is based on distance.

3. System Preliminaries

This paper considers a sensor network consists of N number of sensor nodes randomly deployed over a vast area. The sensors continuously sensed the area and send the sensed data to the sink located outside the area, through intermediate sensors. Each sensor node can operate either as a cluster member to sense the environment and send the data to a CH or as a CH to collect data compress it and send it to the sink. Apart from this functionality, a CH can work as a relay node. In addition, the following assumptions have been taken.

- Sensors and the sink all are stationary after deployment.
- All the sensors are homogeneous and have the same amount of initial energy.
- Sensors are left unattended after deployment. Therefore, the battery recharge is not possible.
- Each sensor is uniquely identified by an identifier (node ID).
- Sensors have the capability to vary the amount of transmission power depending on the distance of the receiving node.
- The distance between the sensors can be calculated based on received signal strength if the transmitting power is known.
- The radio links are symmetric. Thus, the communication between any two nodes required the same transmission power.

Currently, the research is more focused on the area of low energy radios. In this paper, Heinzelman et al. [13] apply the first-order radio model to model the energy dissipation. If the distance between the receiver and the transmitter is lower than a threshold value d_0, the free space model (d^2 power loss) is used. Otherwise, the multipath fading channel model (d^4 power loss) is utilised. The energy required to transmit and receive an l-bit packet over a distance d is given by Eqs. (1) and (2) as follows:

$$E_{Tx}(l,d) = lE_{elec} + l\grave{o}d^{\alpha} = \begin{cases} lE_{elec} + l\grave{o}_{fs}d^2, d < d_0 \\ lE_{elec} + l\grave{o}_{mp}d^4, d \geq d_0 \end{cases} \tag{1}$$

$$E_{Rx}(l) = lE_{elec} \tag{2}$$

where E_{elec} energy is required to run the transceiver, it depends on factors like digital coding and modulation and $\epsilon_{fs}d^2$ or $\epsilon_{mp}d^4$ is the amplifier energy that depends on the transmission distance and an acceptable bit error rate. The threshold value d_0 can be obtained from Eq. (3) as follows:

$$d_0 = \sqrt{\grave{o}_{fs}/\grave{o}_{mp}} \tag{3}$$

It is assumed that the sensed information is highly co-related, therefore the CH always aggregate the data gathered from the member node and compressed it into a single packet. In some proposed methodology, the relay node can also take care of the data aggregation. This is not feasible because the co-relation factor is quite negligible for data from different clusters. Therefore, in the proposed approach, the relay node does not aggregate the relay packets. In the proposed methodology, it is assumed that E_{DA} *(nJ/*bit/signal*)* amount of energy is consumed by the CH for data aggregation.

4. Proposed Approach

The on-demand clustering algorithms proposed earlier show excellent improvement over round by round based clustering approach. However, the on-demand clustering approach proposed so far uses equal cluster range for CHs. In a multi-hop inter-cluster communication scenario, the CHs near the sink is overburdened with relay traffic and lose a significant amount of energy in relaying data, thus the CHs near the sink trigger the clustering process frequently.

Hence, based on the above observation, the proposed algorithm constructs a more balanced clustering scheme by considering nodes remaining energy while selecting the CHs and assigning different cluster range to CHs. The clustering process of the proposed approach is performed on a demand basis. The CHs are selected based on local information. A node waits a certain amount of time before declaring itself as CH where the wait time is inversely proportional to nodes remaining energy. A FIS is used for calculation of cluster range of CHs. The overall process of the proposed approach is depicted in Fig. 1. Although there are approaches exist on on-demand clustering, and FIS based unequal clustering, however, the proposed approach is first of its kind to use on-demand clustering along with FIS based inequality in clustering to ensure better load distribution. Moreover, the centrality is a vital parameter in deciding the load, which is overlooked by these FIS, based approaches proposed so far.

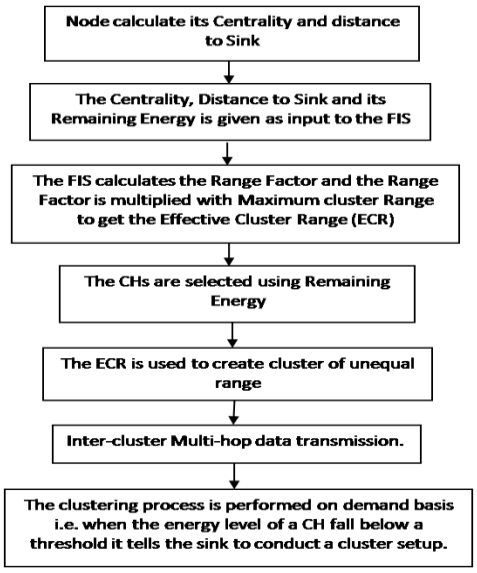


Fig. 1. Overview of proposed approach.

The operation of the proposed methodology is divided into rounds. For each round, the proposed approach consists of three stages: neighbourhood discovery phase, cluster set-up phase, and the steady-state phase. However, the neighbourhood discovery phase occurs once at the time of network deployment before the actual operation of the proposed approach begins. Therefore, it can be said that each round consists of cluster set-up phase and the steady-state phase. Moreover, in the proposed approach, the cluster setup phase is followed by multiple steady-state phases.

4.1. Neighbourhood discovery phase

The algorithm starts with the neighbourhood discovery phase, in which, the sink broadcast a Hello message. On receiving this Hello message, a node can calculate its distance from the sink [23]. Receiving node of the Hello message, broadcast a Hello Reply message consists of sender id, within a range R_{max}. Where R_{max} is the maximum cluster range. Receiving nodes of the Hello_Reply message add the sender as its neighbour. Whenever any node has remaining energy below a given threshold, it will broadcast itself as dead by sending a Dead message. The receiving nodes of dead message update their neighbourhood information.

4.2. Cluster set-up phase

The cluster setup phase of the proposed approach consists of:

- Selection of candidate CHs.
- Assigning appropriate cluster range to candidate CHs.
- Selection of the CHs from the set of candidate CHs.
- Assigning non-CH nodes to clusters.

After the end of the neighbourhood discovery phase, each node waits for a $Wait_{Time}$ before broadcast the Candidate_CH message. The calculation of $Wait_{Time}$ is given by Eq. (4) as follows:

$$Wait_{Time} = \frac{1}{Remaining\ Energy} \tag{4}$$

When the Wait Time is over, the candidate CH broadcast a Candidate_CH message in its Effective Cluster Range (ECR). The calculation of ECR is outlined in the subsequent section.

The candidates CHs are the set of nodes, which have sent Candidate_CH message and after that either they have not received any Candidate_CH message or their remaining energy is higher than the neighbours. It may also be possible that a node has less remaining energy value sent Candidate_CH message, in that case, the node will send a Quit_Election message and decides to be a cluster member. This competition guaranteed that only the node with the best remaining energy will be the CH and there will be no other CH in its ECR. The remaining candidate CHs sends a Final_CH message in twice its clustering range. Twice the range of this constraint is required to inform the neighbour CH about its election as CH. The Final_CH message consists of node ID and distance to sink. On receiving a Final_CH message, the CHs maintain a table of its neighbour CHs. The entry in that table is neighbour node ID and its distance to sink. If every non-CH node

receives multiple final CH messages, it will join the shortest distant CH by sending a Membership message. Now, each CH prepares a Time Division Multiple Access (TDMA) schedule telling each member node when to transmit for which, it received Membership messages. This TDMA schedule is broadcast back to the member nodes in the cluster.

4.3. Fuzzy-logic based cluster range calculation

To handle ambiguity in cluster range selection, in this paper the FIS has been used. The proposed approach is based on the Mamdani method of fuzzy inference technique [24]. Figure 2 gives a detailed outline of the FIS used in the proposed work. At first, the inputs are transferred into a linguistic term called fuzzification. This process is accomplished with the help of membership functions. Then the fuzzy inference unit uses the fuzzy "if-then" rules to map the fuzzy inputs to fuzzy outputs. Finally, the fuzzy outputs are defuzzified to crisp values, called defuzzification.

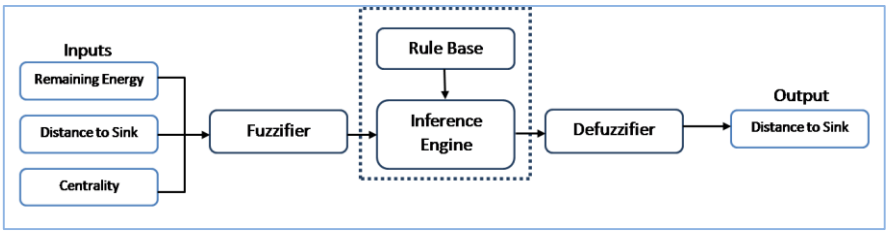


Fig. 2. Outline of the FIS used in the proposed work.

The input variables of the FIS are Remaining Energy, Distance to the sink and Centrality. The only output of the FIS is Range Factor. The input variables are defined as follows.

Remaining energy: Energy level remaining of the node.

Distance to sink: Nodes distance (Euclidean distance) to the sink.

Centrality: A value based on nodes distance from the neighbour's proportion to network dimension.

The value of the actual inputs can be different for different network setup. Therefore, in this paper, the values of the inputs are in a range of 0 to 1. The inputs parameters of different values are scaled to fit in this range. The input variable Remaining Energy is calculated using Eq. (5) as follows:

$$Remaining\ Energy = \frac{E_{current}}{E_{initial}} \tag{5}$$

where, $E_{initial}$ is the initial energy of the node and $E_{current}$ is the current energy of the node.

The input variable Distance to sink is calculated using Eq. (6) as follows:

$$Distance\ to\ Sink = \frac{D_i}{D_{\max}} \tag{6}$$

where, D_i is node's distance from the sink and D_{max} is the distance of the furthermost node from the sink.

The input variable centrality is calculated using Eq. (7) as follows:

$$Centrality = \frac{\sqrt{\left(\sum_{j \in N_{Nbr}(i)} dist^2(i,j)\right) \Big/ N_{Nbr}(i)}}{NetworkDimension} \tag{7}$$

where, N_{Nbr} signifies neighbours of the node and $dist^2(i, j)$ is the square of the distance between node i and j.

The first input fuzzy set is the remaining energy, Fig. 3 illustrates membership functions of input variable remaining energy. The fuzzy sets in the form of linguistic variables include low, medium and high. The second input fuzzy set is

Distance to sink of the node, Fig. 4 illustrates membership functions of input variable distance to sink. The fuzzy sets in the form of linguistic variables include near, not too far and far.

The third input fuzzy set is node centrality, Fig. 5 illustrates membership functions of input variable centrality. The fuzzy sets in the form of linguistic variables include close, acceptable and far. The only fuzzy output variable of the FIS is the range factor, Fig. 6 illustrates membership functions of output variable range factor. There are five linguistic variables for the output fuzzy set fitness; they are very small, small, medium, large, and very large.

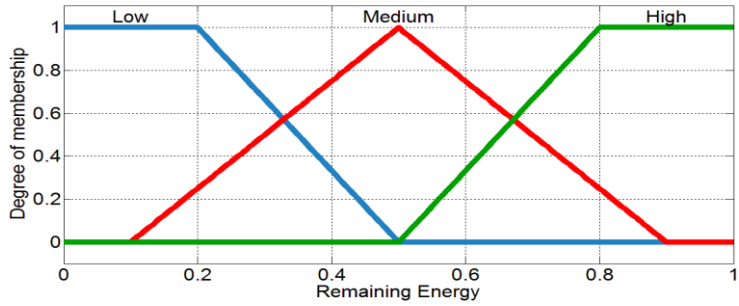


Fig. 3. Membership functions of input variable remaining energy.

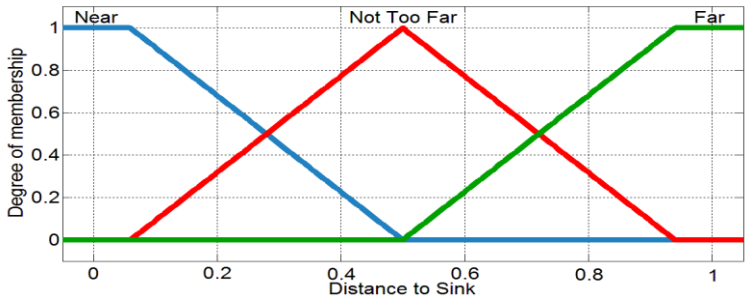


Fig. 4. Membership functions for input variable distance to sink.

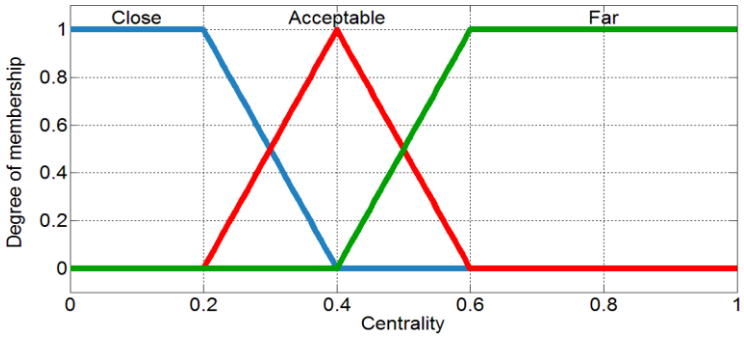


Fig. 5. Membership functions for input variable centrality.

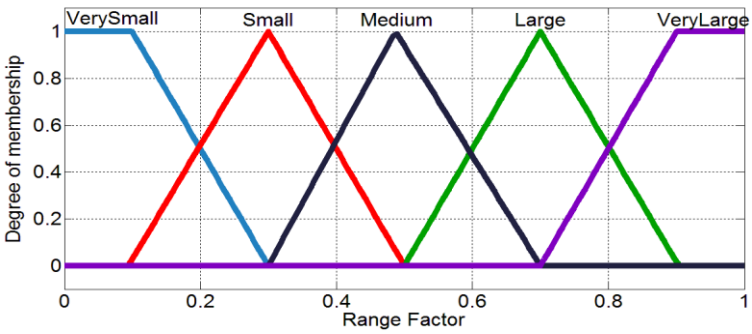


Fig. 6. Membership functions for output variable range factor.

Several of candidate memberships functions are tested to find the best fitting functions for all the inputs and output variables and based on the result these membership functions are chosen.

For this work, mostly triangular membership function has been used because triangular membership functions are quite simple, a small amount of data is required to define it, easy to modify and have computational efficiency compared to other membership functions [25].

Fuzzy inference unit uses the fuzzy "if-then" rules for the calculation of output. There are twenty-seven fuzzy mapping rules based on the combination of different linguistic variables specified in Table 1. By applying these fuzzy "if-then" mapping rules on the inputs, the fuzzy output Range Factor has been generated. The output fuzzy variable has no use unless defuzzified to a single crisp value. In this paper, the Centre of Gravity (COG) method [26] has been used for the defuzzification. The method is given by Eq. (8) as follows.

$$Z^* = \int U_i(Z) Z dZ \Big/ \int U_i(Z) dZ \tag{8}$$

where Z^* is the defuzzified output, $U_i(Z)$ the aggregated membership function and Z is the output variable.

The whole logic of the FIS is required to be embedded in the sensor nodes. The processing unit of the sensor makes use of the embedded FIS to decide the Range

Factor. In addition, as this approach is a decentralized one, the FIS capability needs to be implemented by every sensor nodes. Usually, fuzzy rules are either generated from experimental results or based on heuristics. For this work, the heuristic based fuzzy rule generation method has been used.

Range Factor is the variable that influences the cluster range of the candidate CHs. The ECR of the candidate CHs is calculated using Eq. (9) as follows:

$$Effective\ Cluster\ Range(ECR) = Range\ Factor \times R_{\max} \tag{9}$$

Table 1. Fuzzy rule base.

SI No.	Remaining energy	Distance to sink	Centrality	Range factor
1	Low	Near	Far	Very small
2	Low	Near	Acceptable	Very small
3	Low	Near	Close	Small
4	Low	Not too far	Far	Very small
5	Low	Not too far	Acceptable	Small
6	Low	Not Too Far	Close	Small
7	Low	Far	Far	Small
8	Low	Far	Acceptable	Medium
9	Low	Far	Close	Medium
10	Medium	Near	Far	Small
11	Medium	Near	Acceptable	Medium
12	Medium	Near	Close	Medium
13	Medium	Not too far	Far	Large
14	Medium	Not too far	Acceptable	Medium
15	Medium	Not too far	Close	Small
16	Medium	Far	Far	Large
17	Medium	Far	Acceptable	Medium
18	Medium	Far	Close	Large
19	High	Near	Far	Medium
20	High	Near	Acceptable	Large
21	High	Near	Close	Large
22	High	Not too far	Far	Medium
23	High	Not too far	Acceptable	Large
24	High	Not too far	Close	Large
25	High	Far	Far	Large
26	High	Far	Acceptable	Very large
27	High	Far	Close	Very large

4.4. Multi-hop inter-cluster communication

The data transmission begins after the CHs are selected and TDMA schedule is fixed. Every member node sends its data according to the predefined TDMA schedule and goes to sleep state. However, the CH node must keep its receiver on, to receive data from the member nodes of its cluster. In this work, the authors use a multi-hop inter-cluster communication model for communication between CH and the sink. Therefore, after receiving all the data from member nodes, the CH compresses the data and passes it to the next CH or to the sink directly (distance to sink $\leq R_{max}$ or no such CH exists). In most of the applications of the wireless sensor network, the minimum delay is a prerequisite. Therefore, the selection of relay CH in this work is put based on the distance to sink. The CH selects the neighbour CH

with the minimum distance to sink from recorded neighbouring CH information during the cluster setup phase. This is given by Eq. (10) as follows:

$$CH_{Relay} = \begin{cases} \min\left(CH_{Nbr}.Distance\ to\ Sink\right), & Distance\ to\ Sink > R_{\max} \\ Sink, & Distance\ to\ Sink \le R_{\max} \\ Sink, & CH_{Nbr} = \varnothing \end{cases} \tag{10}$$

where CH_{Relay} is the selected next-hop CH, CH_{Nbr} is the set of neighbour CH and CH_{Nbr}.Distance to sink is the distance of neighbour CH from the sink.

The clustering process of the proposed approach is on a demand basis. Therefore, after every steady-state phase, the CH compares its remaining energy with the predetermined threshold, if it has left with energy less than the threshold, it reports to the sink to conduct a cluster setup.

4.5. On-demand clustering

Most of the clustering algorithm proposed earlier select CH in every round. Their action is justified, as it selects the CHs with the best resources. However, the problem is; a clustering process required many control information exchanges. Therefore, in this paper, Taheri et al. [7] proposed an on-demand clustering process, in which similar to the clustering phase is triggered by the sink when the remaining energy of a CH goes beyond a given threshold, and this process is described using a flowchart in Fig. 7.

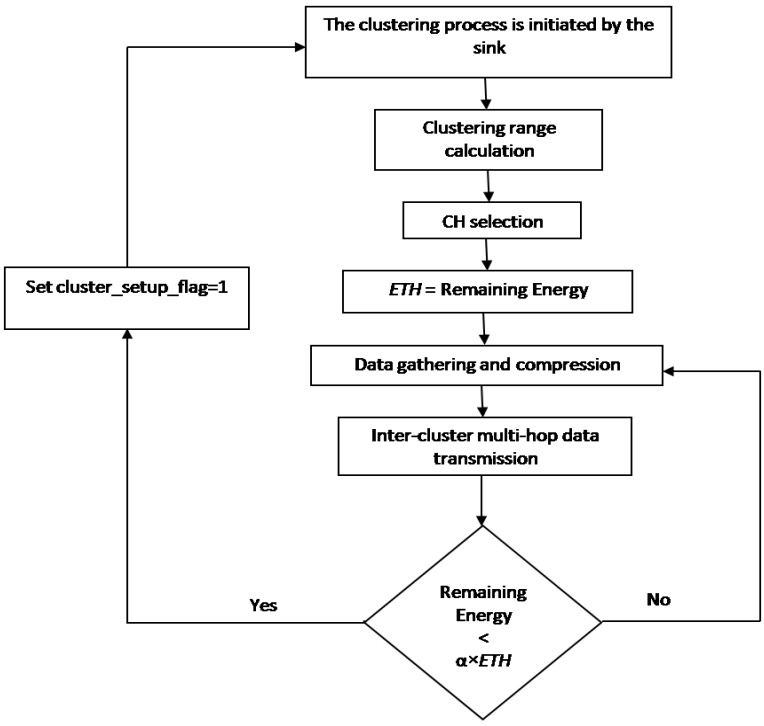


Fig. 7. Flowchart of on-demand clustering process.

After the end of each cluster setup phase, each selected CH save its energy status in a variable E_{Re} and stored it in the memory. At the end of every steady-state phase, if the remaining energy of any of the CHs falls below $\alpha \times E_{Re}$ (α floating point number with range $0 \leq \alpha \leq 1$), it set the cluster_setup_flag in the packet as 1 and send it to the sink. On receiving a packet if the sink found that cluster_setup_flag set as 1 it broadcast a pulse signal that informs the nodes to conduct a cluster setup. The on-demand clustering process mitigates the control packet overhead required by consecutive cluster setup phase consequently, decrease the overall energy consumption.

5. Result and Discussion

In this section, the results of the simulation experiments are presented to illustrate the effectiveness of the proposed method. For simplicity, in this work, an ideal Media Access Control layer and an error-free communication link are used. Moreover, energy is consumed whenever a sensor sends or receives data or performs data aggregation.

To demonstrate the effectiveness of the proposed algorithm, it is compared with ECPF and LEACH. All these algorithms are implemented and simulated for different network configuration. The simulations are performed using MATLAB. All experiments are conducted on an Intel Core i5 processor server running the Windows 7 operating system. Every scenario is simulated multiple times and the average results are drawn from these experimentations are presented over here.

In order to evaluate the performance of the proposed approach, four different network configurations are used. The details of the network configurations are given in Table 2 and for a better comparison, we used the almost same simulation parameter that has been used to simulate ECPF. The details of the parameters used in the simulation with their values are given in Table 3.

For this work, Heinzelman et al. [13] stated that the first-order radio model is used to model the energy dissipation. The simulation is performed with 100, 200, 300 and 400 set of nodes. The metrics used for comparison are the lifetime of the network in terms of rounds, i.e., the First Node Dies (FND), half of the Nodes Alive (HNA), the Last Node Dies (LND) [27] and based on the energy spent per round.

Table 2. Network configurations used for simulation.

Configuration no.	Parameters	Value
Configuration-I	Network dimension	100 m, 100 m
	Sink position	50 m, 50 m
	R_{max}	70 m
Configuration-II	Network dimension	100 m, 100 m
	Sink position	50 m, 175 m
	R_{max}	70 m
Configuration-III	Network dimension	200 m, 200 m
	Sink position	100 m, 100 m
	R_{max}	110 m
Configuration-IV	Network dimension	200 m, 200 m
	Sink position	100 m, 275 m
	R_{max}	110 m

Table 3. Parameters used for simulation.

Parameter	Value
Initial energy per node	2 J
ϵ_{fs}	10 pJ/bit/m^2
ϵ_{mp}	0.0013 pJ/bit/m^4
E_{elec}	50 nJ/bit
E_{da}	5 nJ/bit/signal
D_0	87 m
Idle power	13.5 mW
Sleep power	15 μW
Round time	20 s
Data packet size	100 byte
Control packet size	25 byte
Aggregation ratio	10%

In each round of the simulation, first, the CHs are elected than the cluster is formed, after that, every non-CH node sends data to the CH and the CH aggregate the data and forwards it to the next CH or to the sink.

As stated by Bagci and Yazici [11], the aggregation ratio for the simulation is set to 10% using the same mechanism and the value of α is set to 0.8 as stated by Taheri et al [7]. These values are also found suitable by running simulations according to network configurations and simulation parameters as stated in Tables 2 and 3 respectively.

The results are plotted based on average energy consumption per round and lifetime metrics FND, HNA and LND. The operation of LEACH is conducted in rounds, wherein every round the clusters is formed, data is collected and transmitted to the sink, i.e., one cluster setup phase is followed by one steady-state phase.

However, in case of ECPF and the proposed approach one cluster setup phase is followed by multiple steady state phases. Therefore, for simplicity one steady state phase is considered as one round of network operation.

5.1. Network configuration-I

In this network configuration nodes are deployed over a 100 m × 100 m area and the sink is placed at the centre of the network. The aim of this configuration is to assess the behaviour of the clustering approach when the sink is placed at the centre of the network. The details of this network configuration along with the other three are given in Table 2.

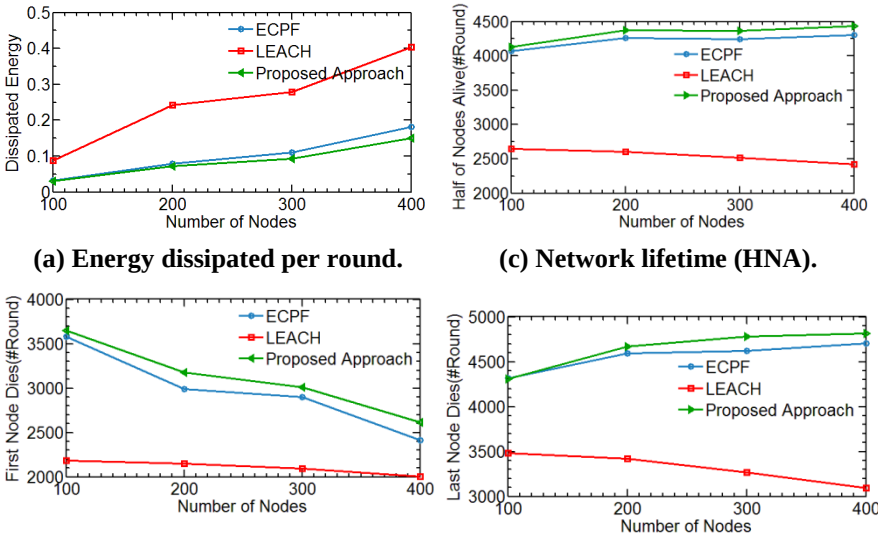
The maximum competition radius is set to 70 m as given in Table 2. This is the maximum competition radius values found optimal for this configuration. The results of the simulation for this configuration are depicted in Fig. 8. Figure 8(a) represents the average energy consumption per round.

It can be observed that the performance of the proposed algorithm is better than LEACH and ECPF; moreover, the result is getting better with an increase in a number of nodes. The comparison of the network lifetime of the proposed approach with ECPF and LEACH is demonstrated in Figs. 8(b) to (d).

The comparison of the network lifetime is based on three metrics namely FND, HNA, and LND. It can be observed that in contrast to HNA and LND metrics the

performance of all the approaches in FND metric degraded with an increase in a number of nodes. This happens due to the employed shortest distance based multi-hop routing approach overused the CHs in the shortest route to the sink. However, it can be observed from the figures that the performance of the proposed approach is well ahead of ECPF and LEACH.

The probabilistic selection of CHs in LEACH creates an unbalanced energy consumption structure resulting in poor performance. The fuzzy-logic based unequal clustering of the proposed approach distributes the network load optimally over the selected CHs. Thus, it solves the immature triggering of the clustering process of ECPF. As the number of clustering process decreased, consequently the average energy consumption per round decreased.



(a) Energy dissipated per round. **(c) Network lifetime (HNA).**

(b) Network lifetime (FND). **(d) Network lifetime (LND).**

Fig. 8. Results of the simulation for Network Configuration-I.

5.2. Network configuration-II

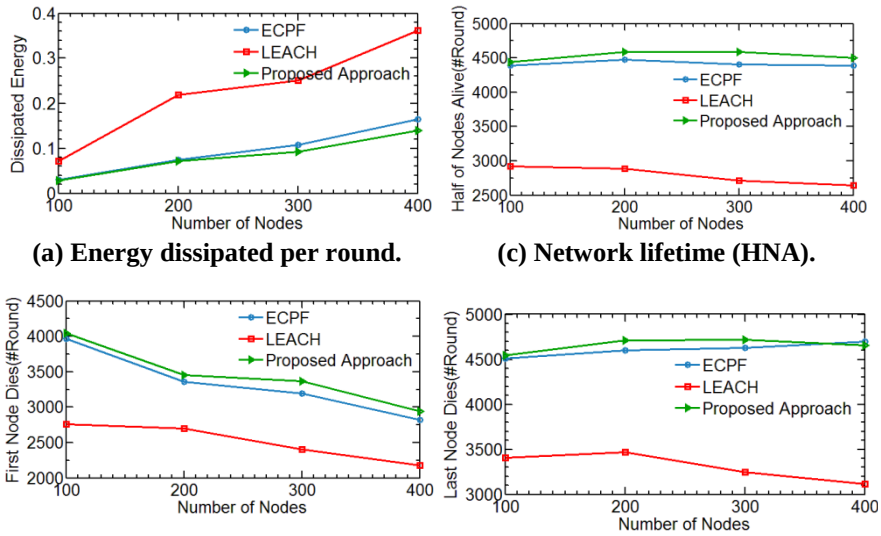
In this network configuration, the deployment of nodes is the same as of configuration-I but the sink is placed outside the network. The aim of this configuration is to assess the behaviour of the clustering approach when the sink is placed outside the network.

The details of this network configuration are given in Table 2. Figure 9 illustrates the results of the simulation for this configuration. The average energy consumption per round is shown in Fig. 9(a). Again the performance of the proposed algorithm is better than LEACH and ECPF and it is getting better with an increase in a number of nodes.

Figures 9(b) to (d) represent the network lifetime metrics FND, HNA and LND respectively. In this scenario, also the performance in FND metric is degraded with an increase in a number of nodes. This happens due to the same (overuse of shortest route to the sink) fact as discussed in scenario-I. Moreover,

like the previous scenario the direct mode of communication of LEACH resulting in the poorest performance.

Though ECPF uses a multi-hop approach for inter-cluster data routing the unbalanced energy consumption structure due to equality in clustering cost it with more number of elections. Therefore, again the performance of the proposed approach is much better than ECPF and LEACH.



(a) Energy dissipated per round. **(c) Network lifetime (HNA).**

(b) Network lifetime (FND). **(d) Network lifetime (LND).**

Fig. 9. Results of simulation for Network Configuration-II.

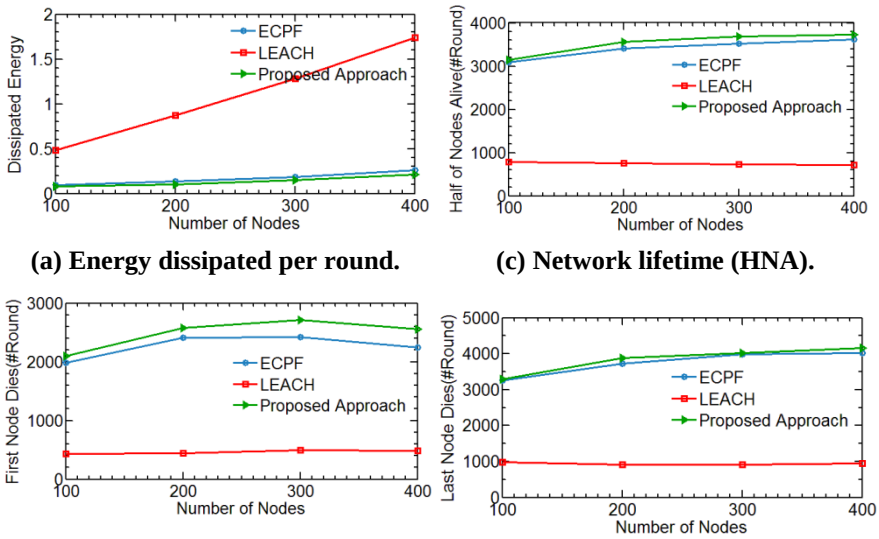
5.3. Network configuration-III

In this network configuration nodes are deployed over a 200 m × 200 m area and the sink is placed in the middle of the network. The aim of this configuration is to assess the behaviour of the clustering approach when the network is scaled to a larger area.

Therefore, the central idea behind this configuration is to exploit the dimension scalability of the clustering approaches. The maximum competition radius is set to 110 m, as given in Table 2. All the results of this simulation are displayed in Fig. 10.

Figure 10(a) is represented based on average energy consumption per round. The performance of the proposed algorithm is much better than LEACH compared to first two configurations.

The proposed algorithm shows a better result with an increase in the number of nodes when compared with ECPF. The network lifetime metrics FND, HNA and LND are displayed in Figs. 10(b) to (d) respectively.



(a) Energy dissipated per round. **(c) Network lifetime (HNA).**

(b) Network lifetime (FND). **(d) Network lifetime (LND).**

Fig. 10. Results of simulation for Network Configuration-III.

However, in this scenario, the performance in **the** FND metric gets better with an increase in the number of nodes. This happens due to the increased in the area of deployment. The increase in deployment area pushes the network to form a large set of clusters. These clusters provide distinct path option in multi-hop routing thus minimizing the overuse of some routing paths.

However, similar to other scenarios the performance of LEACH is significantly lower in this case also. The reason behind this poor performance is the same as mentioned in the previous configurations. However, it can be observed that the proposed algorithm show a better result than ECPF in the FND metrics, also the performance of the proposed approach is quite good compared to ECPF in the other two metrics. It is due to the same fact as discussed in the former configurations. This justifies that the proposed approach is scalable with respect to the network dimension.

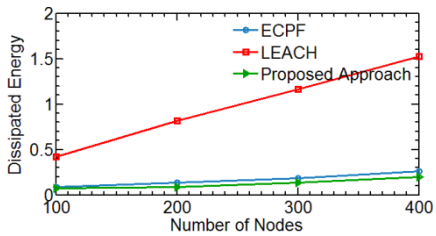
5.4. **Network configuration-IV**

This network configuration is similar to configuration-III; the only change is the sink is placed outside the network. The aim of this configuration is to assess the effect of the sink when placed outside the network in large scale dimension of the network. Figure 11 summarises the overall performance of all the approaches for this configuration. Figure 11(a) displayed a line chart representing the average energy consumption per round. It can be observed from the figure that like the other three configurations, the performance of the proposed algorithm is better than LEACH and ECPF. However, like other three configurations, the proposed approach shows a considerable level of improvement over ECPF in all three matrices. In this configuration, the difference in FND metrics is even better compared to ECPF, which can be analysed in Fig. 11(b).

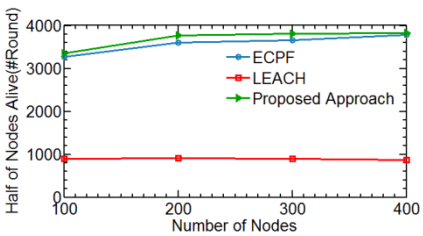
In Fig. 11(c), the HNA metrics with respect to a number of rounds is displayed and in Fig. 11(d), the LND metrics with respect to a number of rounds is displayed. However, like the previous one in this scenario also the performance in FND metric is not degraded with an increase in a number of nodes. This happens due to the same fact as discussed in the previous scenario. Therefore, like preceding configurations in this case also the proposed algorithm did well to outperform the compared approaches.

By comparing all the four network configurations, it can be summarised that the performance of the proposed approach is getting better with an increase in a number of nodes; moreover, it gives a better result with a large number of nodes, which, justifies the scalability of the proposed approach in respect to the number of nodes. By comparing network configurations I, II with network configurations III, IV, it can be observed that the proposed approach is not affected by the change in dimension of the network, which justifies the proposed approach is scalable in respect to network dimension.

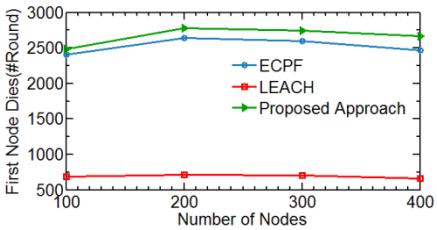
By comparing network configurations I, III with network configurations II, IV it can be observed that the proposed approach is less affected by the change in the position of the sink, which justifies the proposed approach is scalable in respect to the sink position.



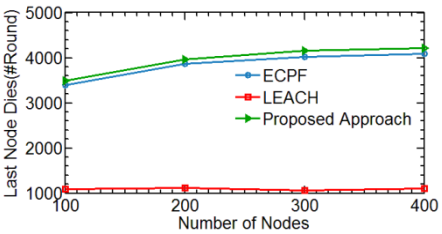
(a) Energy dissipated per round.



(c) Network lifetime (HNA).



(b) Network lifetime (FND).



(d) Network lifetime (LND).

Fig. 11. Results of simulation for Network Configuration-IV.

6. Conclusion

Better lifetime is a design goal that every clustering approach strives to accomplish. This paper presents a fuzzy-logic based unequal clustering approach along with an on-demand based clustering to efficiently improve the lifetime of the WSN. The fuzzy-logic based unequal cluster range approach efficiently assigns different cluster range to CHs based on its remaining energy, distance to sink and centrality,

thus create a stable energy consumption structure, with on-demand clustering minimizes the number of CH selection.

All the approaches are thoroughly examined by doing extensive simulations. According to the simulation result, the proposed approach outperforms ECPF and LEACH in every aspect. The results showed that the proposed method is scalable and improves the network lifetime significantly irrespective of a number of nodes, network dimension, and position of the sink.

Nomenclatures

CH_{Nbr}	Set of neighbour cluster heads of a cluster head
CH_{Relay}	Selected next-hop cluster head for relaying packets
D_i	Distance of ith node from the sink
D_{max}	Maximum possible distance between any node and the sink
$dist^2(i,j)$	Square of the distance between node i and j
d_0	Threshold distance
$E_{current}$	Remaining energy of a node
E_{DA}	Energy required for data aggregation
E_{elec}	Energy is required to run the transceiver
$E_{initial}$	Initial energy of a node
E_{Re}	Remaining energy of a node at the end of cluster setup phase
E_{Rx}	Energy required to receive a packet
E_{Tx}	Energy required to transmit a packet
N_{Nbr}	Neighbour set of a node
R_{max}	Maximum clustering range
$U_i(Z)$	Aggregated fuzzy membership function
$Wait_{Time}$	Waiting time of nodes before broadcasting candidate message
Z^*	Defuzzified output value
Z	Fuzzy output variable

Greek Symbols

α	A floating point value between 0 and 1
ϵ_{fs}	Energy loss in free space channel
ϵ_{mp}	Energy loss in multi-path fading channel

Abbreviations

CH	Cluster Head
COG	Centre of Gravity
ECPF	Energy-aware distributed dynamic Clustering Protocol
ECR	Effective Cluster Range
FIS	Fuzzy Inference System
FND	First Node Dies
HEED	Hybrid Energy-Efficient Distributed
HNA	Half of the Nodes Alive
LEACH	Low-Energy Adaptive Clustering Hierarchy
LND	Last Node Dies
TDMA	Time Division Multiple Access
UCR	Unequal Cluster-based Routing
WSN	Wireless Sensor Network

References

1. Akyildiz, I.F.; Su, W.; Sankarasubramaniam, Y.; and Cayirci, E. (2002). A survey on sensor networks. *IEEE Communications Magazine*, 40(8), 102-114.
2. Akyildiz, I.F.; Su, W.; Sankarasubramaniam, Y.; and Cayirci, E. (2002). Wireless sensor networks: A survey. *Computer Networks*, 38(4), 393-422.
3. SenthilKumar, K.; and Amutha, R. (2015). Energy-efficient cooperative communication in wireless sensor networks using turbo codes. *Australian Journal of Electrical and Electronics Engineering*, 12(4), 293-300.
4. Kumar, K.S.; and Amutha, R. (2016). An algorithm for energy efficient cooperative communication in wireless sensor networks. *KSII Transactions on Internet and Information Systems*, 10(7), 3080-3099.
5. Pal, V.; Singh, G.; and Yadav, R.P. (2012). SCHS: Smart cluster head selection scheme for clustering algorithms in wireless sensor networks. *Wireless Sensor Network*, 4(11), 273-280.
6. Wei, C.; Yang, J.; Gao, Y.; and Zhang, Z. (2011). Cluster-based routing protocols in wireless sensor networks: A survey. *Proceedings of the International Conference on Computer Science and Network Technology*. Harbin, China, 1659-1663.
7. Taheri, H.; Neamatollahi, P.; Younis, O.M.; Naghibzadeh, S.; and Yaghmaee, M.H. (2012). An energy-aware distributed clustering protocol in wireless sensor networks using fuzzy logic. *Ad Hoc Networks*, 10(7), 1469-1481.
8. Li, C.; Ye, M.; Chen, G.; and Wu, J. (2005). An energy-efficient unequal clustering mechanism for wireless sensor networks. *Proceedings of the International Conference on Mobile Adhoc and Sensor Systems Conference*. Washington, D.C., United States of America, 597- 604.
9. Chen, G.; Li, C.; Ye, M.; and Wu, J. (2009). An unequal cluster-based routing protocol in wireless sensor networks. *Wireless Networks*, 15(2), 193-207.
10. Gupta, I.; Riordan, D.; and Sampalli, S. (2005). Cluster-head election using fuzzy logic for wireless sensor networks. *Proceedings of the 3rd Annual Communication Networks and Services Research Conference (CNSR'05)*. Halifax, Nova Scotia, Canada, 255-260.
11. Bagci, H.; and Yazici, A. (2013). An energy aware fuzzy approach to unequal clustering in wireless sensor networks. *Applied Soft Computing*, 13(4), 1741-1749.
12. Heinzelman, W.R.; Chandrakasan, A.; and Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*. Maui, Hawaii, United States of America, 1-10.
13. Heinzelman, W.B.; Chandrakasan, A.P.; and Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1(4), 660-670.
14. Xiangning, F.; and Yulin, S. (2007). Improvement on LEACH protocol of wireless sensor network. *Proceedings of the International Conference on Sensor Technologies and Applications (SENSORCOMM)*. Valencia, Spain, 260-264.

15. Younis, O.; and Fahmy, S. (2004). HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing*, 3(4), 366-379.
16. Younis, O.; and Fahmy, S. (2004). Distributed clustering in ad-hoc sensor networks: A hybrid, energy-efficient approach. *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies, Volume 1 (INFOCOM).* Hong Kong, China, 640.
17. Soro, S.; and Heinzelman, W.B. (2005). Prolonging the lifetime of wireless sensor networks via unequal clustering. *Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium*. Denver, Colorado, United States of America, 8 pages.
18. Kim, J.-M.; Park, S.-H.; Han, Y.-J.; and Chung, T.-M. (2008). CHEF: Cluster head election mechanism using fuzzy logic in wireless sensor networks. *Proceedings of the 10th International Conference on Advanced Communication* Technology. Gangwon-Do, South Korea, 654-659.
19. Anno, J.; Barolli, L.; Xhafa, F.; and Durresi, A. (2007). A cluster head selection method for wireless sensor networks based on fuzzy logic. *Proceedings of the IEEE Region 10 Conference* (*TENCON)*. Taipei, Taiwan, 1-4.
20. Bagci, H.; and Yazici, A. (2010). An energy aware fuzzy unequal clustering algorithm for wireless sensor networks. *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ)*. Barcelona, Spain, 1-8.
21. Sert, S.A.; Bagci, H.; and Yazici, A. (2015). MOFCA: Multi-objective fuzzy clustering algorithm for wireless sensor networks. *Applied Soft Computing*, 30, 151-165.
22. Adhikary, D.R.D.; and Mallick, D.K. (2015). A fuzzy-logic based relay selection scheme for multi-hop wireless sensor networks. *Proceedings of the 1st International Conference on Next Generation Computing Technologies (NGCT).* Dehradun, India, 285-290.
23. Patwari, N.; Ash, J.N.; Kyperountas, S.; Hero, A.O.; Moses, R.L.; and Correal, N.S. (2005). Locating the nodes: Cooperative localization in wireless sensor networks. *IEEE Signal Processing Magazine*, 22(4), 54-69.
24. Wang, L.-X. (1999). *A course in fuzzy systems* (6th ed.). New Jersey, United States of America: Prentice-Hall Press.
25. Pedrycz, W. (1994). Why triangular membership functions? *Fuzzy Sets and Systems*, 64(1), 21-30.
26. Runkler, T.A. (1997). Selection of appropriate defuzzification methods using application specific properties. *IEEE Transactions on Fuzzy Systems*, 5(1), 72-79.
27. Handy, M.J.; Haase, M.; and Timmermann, D. (2002). Low energy adaptive clustering hierarchy with deterministic cluster-head selection. *Proceedings of the 4th International Workshop on Mobile and Wireless Communications Network.* Stockholm, Sweden, 368-372.