

Energy Efficient Clustering Algorithm for Wireless Sensor Networks using Fuzzy Logic

M. Taheri
Department of graduate studies
Islamic Azad University
Bushehr, Iran

Yousef S. Kavian
Faculty of engineering
Shahid Chamran University
Ahvaz, Iran

ABSTRACT

The network lifetime is an important issue for employing wireless sensor networks in space and extreme environments. This is due to the fact that the sensing node energy is mainly consumed by transmissions. For maximizing the network lifetime in this paper a multi-hop clustering algorithm using the fuzzy logic improvement methods is introduced. The cluster heads nodes selection is based on four descriptors; residual energy as primary parameter, node proximity to its neighbors (centrality), distance to base station and node concentration. The proposed multi-hop communication in cluster nodes and between cluster heads reduces the consumption of energy in the network. The results of proposed algorithm are compared with LEACH, TLCP and EHEED algorithms in MATLAB environment. The three metrics FND (first node die), HND (half node die) and LNA (last node alive) show the efficiency of proposed algorithm for the network lifetime.

KEYWORDS

Sensor networking; fuzzy logic; lifetime; multi-hop communication; clustering

1. INTRODUCTION

Under the emerging rapid developments of new science and industries the integration of smart advanced distributed sensing technologies is becoming a prerequisite to help with surveillance, monitor, and controlling an environment and associated processes. There are, however, many tradeoffs among parameters including energy, delay, accuracy and architecture types but the most important one is the energy, which is cause to prolonging network lifetime [1]. Network sensors may be implemented in remote field and are not reachable easily. But, they should send their measurement values to base station (BS) until their energy is drained. Then, power unit (batteries) should have a long lifetime due to the unavailability of the sensors. Therefore, network lifetime becomes a critical and very important parameter. investigations show that the network lifetime is strongly depends on used routing protocol Several metrics like FND (first node die), HND (half node die) and LNA (last node alive) are defined in wireless sensor network to justify the operation of network . For instance, FND parameter used in industrial or domestic gas and fire detection devices due to the importance of each sensor nodes, because failure of one node can make catastrophic consequences especially in that application.

2. RELATED WORKS

Smart sensors and network topology are prerequisites to achieve the maximum network lifetime. Here, an accurate routing should be used to not lose any data packet. In this section, several types of routing protocols are explained.

There are two types of routing protocols; flat and hierarchal. Flat protocols distribute routing information to routers that are connected to each other without any organization or segmentation structure between them, while hierarchal protocols often group nodes together by a function into a hierarchy.

One of the flat protocols is DIRECT [2]. In DIRECT any sensor node transmits its data directly to BS without any intermediate node. So it doesn't necessary for any node to processing or retransmitting another node's data. But all of nodes consume a lot of energy to send their data separately in long distance to BS. Thus, the nodes run out of power very soon. GOSSIPING [3] is a hierarchal protocol. Any Sensor selects one of its neighbors randomly as destination and transmits its data to that. The data packets are sending from one node to another until being received at the BS. In GOSSIPING nodes don't need to be compress and processing, but because of random destination selecting, the data pockets may transverse a long path to receive to BS. Another problem of this protocol is delay. The LEACH (Low-Energy Adaptive Clustering Hierarchy) protocol [4] proposes a power-efficient clustering algorithm. This algorithm has two phases: setup phase that runs stochastically with period 't' for selecting CHs. A threshold number T(i) for each node described as follow:

$$T(i) = \begin{cases} \frac{P}{1 - P * (r \bmod \frac{1}{P})} & , \forall i \in G \\ 0 & , \forall i \notin G \end{cases}$$

Where, for node i; P is the cluster head probability, r is the number of the current round and G is the collection of nodes that have not been cluster heads in the last 1/P rounds. Every single node picks a random positive number between zero and one, if the random number was less than threshold and node is not CH for last 1/P rounds, the node becomes CH. The ordinary nodes join to nearest CHs based on power of received message that sent by them to consume less energy and the clusters are formed.

The next phase is interaction phase which data is sent to BS. LEACH uses Time Division Multiple Access (TDMA) as a scheduling mechanism for sending ordinary node's data to CH. By data compressing node's ability, the data packets from several sensors combine to a single packet. This action is done by CHs. LEACH, however, is not suitable for WSN deployed in large regions since it prone to long delays when applied to large sensor networks. Then, authors of [5] proposed a clustering protocol and tried by using genetic algorithm for selecting CHs to have minimum energy consumption while covering network. They used two operators; mutation and one-point crossover. The CHs are

chosen from the population in the current generation based on their fitness. The nodes with higher fitness values are selected as population in the next generation. Residual energy of the nodes is the first parameter for selecting the cluster head and the energy that required a node to send a message toward the BS node is second parameter to select best CHs. Gupta [6] uses fuzzy logic to select a cluster head. In this method, during every round, the sensor that has the most chance is selected as cluster head. Three fuzzy input descriptors are used to calculate the parameters including: residual energy of the nodes, centrality of the nodes, and concentration. These parameters send to base station by nodes and it determines cluster heads. HEED (Hybrid Energy-Efficient Distributed clustering) [7], in this algorithm authors present a distributed clustering to achieve a hybrid of energy and communication cost. The HEED protocol calculates the number of iterations, independent of the network diameter. Each node individually decides to become CH or not by calculating the chance including its residual energy as follows:

$$CH_{prob} = C_{prob} \times (E_{residual} - E_{max})$$

TLCP [8] proposes a two-level cluster protocol for minimizing the energy consumption based on LEACH. In here since one node is responsible for sending all network node's data, the energy of this nodes fall suddenly comparing others in each round, on the other hand by failing this node all network's data has been lost.

EHEED (Extended HEED) [9] is the multi-hop technique that is integrated in HEED. CHs are probabilistically selected by their residual energy and an initial rate of cluster head, ordinary nodes join to them by 'minimum degree cost' or 'maximum degree cost' or 'AMRP' based on the network's application, 'Minimum degree cost' used to distributing cluster head load, 'maximum degree cost' for create dense clusters and 'AMRP' that denotes the average power required by nodes to communicate with the CH is a reasonable solution. Inter-cluster communications depends on the network application. CHs may directly communicate with BS or they can communicate with one another to aggregate their information via multiple hops.

3. SYSTEM MODEL

Hierarchical clustering techniques are very useful to reduce energy consumption [10] in wireless sensor networks. Clustering is used in the applications for load balancing and using efficient resources because of its scalability characteristics for thousands of nodes and it is applicable for data aggregation. Also Clustering is helpful in routing protocols [11]. Sensor network clustering can reduce consuming energy of communication by sending whole cluster's data within cluster heads after aggregating data, and it's not necessary for ordinary nodes to communicate directly with BS.

In this paper, as tried to reduce energy consuming in wireless network's nodes or prolong the network lifetime, a multi-hop routing is used to transmit node's data to BS, because it is obviously that most of node's energy dissipated for communication not for processing or sensing environment [12].

The algorithm consists of 4 steps. Step1 explains the fuzzy logic of selecting cluster head nodes among all nodes. In step 2 the cluster forming is described. Step 3 discusses about cluster setup and dividing cluster into several zones. Step 4 explains the communication between CHs with another and with BS.

3.1 First Step (Fuzzy Logic)

System dynamic behavior is characterized by a set of linguistic fuzzy rules. These rules are based on the knowledge and experience of a human expert within that domain.

The fuzzy logic is more robust even when the parameters and constants are not optimum in values, and can be used when complete knowledge of the plant is absent or when the system parameters can change dynamically. The fusion fuzzy logic with clustering has a lot of advantages. The Fuzzy algorithm based on if-then-else statements provides almost a small portion of memory, with less time of execution, and simple flat mathematics than other approaches. The fuzzy inference system merges the rules to get a fuzzy output. Then the defuzzifier maps the fuzzy output back to a crisp number, which is used for making decisions to a node be CH or not. a simple defuzzification method be used to denoted as center of gravity (COG) in order to produce a crisp output [13].

In this step the best nodes selected to be CHs for gathering and aggregating ordinary node's data then transmit data packets to BS. For selecting the best nodes as CHs, the most effective parameters in fuzzy system should be employed, because CHs are responsible for cluster's data and consume more energy. First parameter is nodes residual energy; CHs consume more energy than ordinary nodes, where the nodes should be selected for more residual energy in current round, and this is the main characteristic for a node to be elected as CH. The second parameter is distance to BS; since we want to use multi-hop clustering, we try to have more CH nearer the BS than farther ones. It causes the cluster radius be smaller than those are farther. The necessity of this action is that the CHs near the BS are intermediate CH and are responsible for receive and transmit the data of farther CHs, whether farther CHs are only responsible for their personal nodes.

The third parameter is concentration; it raises the chance of the nodes in dense area to be CH. The fourth parameter is the centrality. The last two parameters cause the total of distances for communication between CHs and ordinary nodes become minimum. Eventually, the nodes with a greater chance will be CHs. As described four fuzzy descriptor and for each descriptor three triangle membership functions (LOW, MEDIUM, HIGH) is defined (Fig. 1), then 81 fuzzy rule base is produced; and for the output seven triangle membership function (VERYSMALL, SMALL, RARLY SMALL, MEDIUM, RARELY LARGE, LARGE, VERY LARGE) are described (Fig. 2). Table 1 is containing 10 examples for fuzzy system that calculates the chance of any node to being CH.

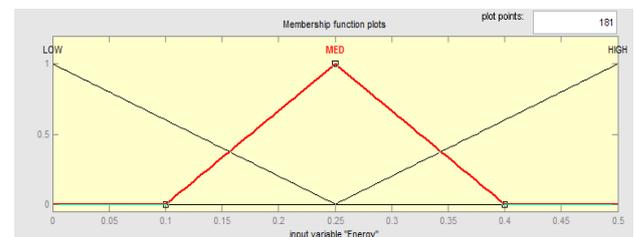


Fig. 1. Membership function for a fuzzy system input (Energy)

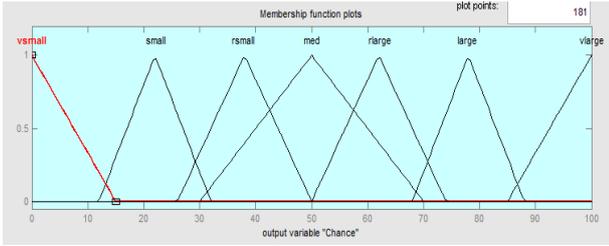


Fig. 2. Membership function for fuzzy system output

TABLE 1. 10 TYPICAL FUZZY SYSTEM RULE BASE

#	Energy	Concentration	Centrality ⁻¹ (Ave-Dis)	Dis to BS	Chance
1	LOW	LOW	HIGH	HIGH	vsmall
2	LOW	MED	MED	LOW	vsmall
3	LOW	LOW	HIGH	MED	small
4	MED	LOW	LOW	HIGH	rsmall
5	MED	LOW	MED	HIGH	rsmall
6	MED	LOW	MED	LOW	med
7	MED	HIGH	LOW	MED	rlarge
8	HIGH	LOW	LOW	HIGH	rlarge
9	HIGH	MED	HIGH	LOW	rlarge
10	HIGH	HIGH	LOW	LOW	vlarge

3.2 Second Step (Cluster Setup)

In this step, formation of clusters is described.

Three type of communication are defined. (i) intra-cluster communication range which should be just enough to cover all the cluster nodes, (ii) CH-BS range, CH can directly communicate with the BS and (iii) inter-cluster communication range [7] that is the communication between CHs. Dynamic clustering causes new cluster heads are elected and new clusters formed per round by calculating the fuzzy chance of each node according to four fuzzy membership function which are explained later.

By the first order radio model, which demonstrates power consumption is proportional to second power of distance, if distance between transmitter and receiver be shorter than d_0 , for longer distances it is proportional to fourth power of distance, whereas d_0 define as follows:

$$d_0 = \sqrt{E_{fs}/E_{mp}}$$

Here E_{fs} is free space and E_{mp} is multi-path fading channel model, E_{elect} is the energy dissipated of transmitter, E_{elec} is expended energy by electronic circuits to serve digital electronics actions such as actuation, sensing, signal emission/reception and d demonstrate distance between transmitter and receiver, k is the number of bits that will be sending, then:

$$E_{TX} = \begin{cases} k \times (E_{elec} + E_{fs} \times d^2), & \text{if } d \leq d_0 \\ k \times (E_{elec} + E_{fs} \times d^4), & \text{if } d \geq d_0 \end{cases}$$

Power consumption for receiving k bits data is expressed by:

$$E_{RX} = k \times E_{elec}$$

With respect to the total power consumed for transmitted and received signal is $u(d)$

$$u(d) = a \times d^\alpha + c$$

Where a is depend on physical environment, d is distance between sender and receiver, c is the energy of data processing in each sensor nodes, α is the power that depend to distance. The parameter's values mentioned are $a=1$, $\alpha = 2$ and, $c = 1000$.

After electing cluster heads, remaining nodes join to nearest cluster head to consume the minimum energy, consequently clusters are formed. To consider energy efficiency for intra-cluster communication the multi-hop protocol is proposed. First, CHs send an advertisement message; all nodes receive several messages from all the CHs and they (nodes) recognize the distance based on received signal power. Then, the nodes select the nearest CH by sending an approval message. It is very effective in energy preservation for nodes to send their data from several smaller paths.

3.3 Third Step(Intra-Cluster communication)

After the formation of clusters, every CH calculates cluster radius by checking power of received signals from all nodes that is equal to distance between CH to farther node. Then, based on cluster radius, CH divides the cluster to several zones and sends the number of zone to related nodes. Intra-cluster communication range, should be just enough to cover all the cluster nodes (Ivan et al., 2001), assuming the distance of sender T to destination R is 'd'. If there is an intermediate node A between T and R, then $|TA| = x$, $|AR| = d - x$ (Fig. 3).

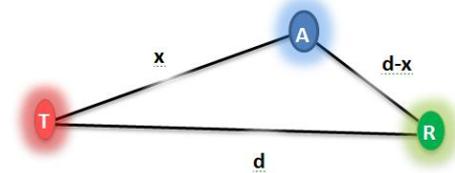


Fig. 3. The message be send from S to D via A

If the following condition is satisfied we have energy efficiency and the best state is when A is in the middle of T and R.

$$d > (c/(a(1 - 2^{1-\alpha})))^{1/\alpha}$$

Then distance could be divided to n parts. n is the nearest integer to the result of following formula.

$$n = d(a(\alpha - 1)/c)^{1/\alpha}$$

Then the minimal power is calculated as:

$$dc(a(\alpha - 1)/c)^{1/\alpha} + dc(a(\alpha - 1)/c)^{(1-\alpha)/\alpha}$$

Like EHEED, the radius of cluster into equally small zones has been divided, nodes based on their distance to related CH getting in a zone if they satisfy the above equations. Any node can easily recognize in which zone it is, by sending the advertisement message from CH to zone 1 nodes and message zone 1 nodes for zone 2 and so on. This message contains node ID, CH-ID, the distance from parent node that is the CH for the first zone and, the Node-Depth (for nodes directly connected to the CH).

If $d > (c/(a(1 - 2^{1-\alpha})))^{1/\alpha}$ communication will be so energy efficient. If distance SD is divided into n parts then the minimal energy is .

$$dc(a(\alpha - 1)/c)^{1/\alpha} + dc(a(\alpha - 1)/c)^{(1-\alpha)/\alpha}$$

With satisfying explained equation, CHs send their data until they get to the BS. For example, if cluster radius is divided into four zones, each %25 of cluster radius will be one zone. Also CH is in the middle of first zone (Fig. 4). This action will save nodes energy because the distance divides to smaller distances.

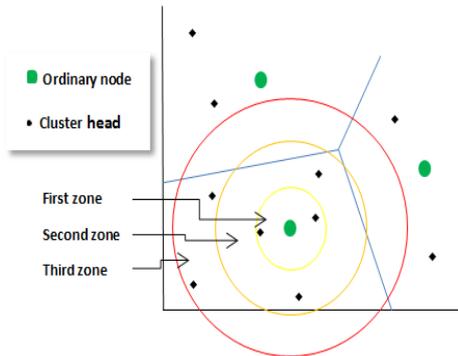


Fig. 4. Dividing zone to three clusters

3.4 Forth Step (Inter-Clustercommunication)

Here node's data is received to related CHs, data packets are produced by CHs with compressing received ordinary node's data and theirs. According to third step description, if CH's data packets want to send via multi-hop path to related base station, the consumption of energy can be reduced. The data of CHs that are farther from BS should convey through nearer CH to BS. There is two types of communications for CHs in this step (Fig. 5): CH-BS; CH can directly communicate with the BS and another one is inter-cluster communication.

BS sends advertisement message, CHs answer the BS to recognize distances and the network is zoned for hierarchical communication. Finally, like intra-cluster communication there are multi-hop communication for inter-cluster communication.

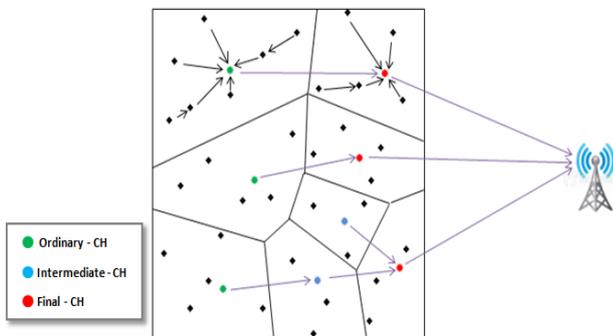


Fig. 5. Algorithm architecture

Like LEACH, in order to reduce collision among joint request messages during third and this step, CSMA (Carrier Sense

Multiple Access) is used as a MAC layer protocol. Nodes, environment and communication features:

1. The network contains one static BS.
2. All nodes in the network are homogenous.
3. All nodes have similar capabilities in processing
4. Nodes are unavailable after deployment.
5. Nodes can detect the signal power.
6. Nodes can adjust their transmission power.
7. Nodes are location-unaware.

3.5 Experimental Results

To evaluate the performance of proposed algorithm, we simulate the algorithm and LEACH, TLCP, EHEED with 150 random nodes in MATLAB.

In this section, the nodes spread into a 100x100 area. All nodes have 0.5j initial energy with maximum 250-meter radio range. A base station is in (200, 50) Cartesian coordinates and the other parameters value mentioned in Table 2.

TABLE 2. PARAMETER VALUES IN SIMULATION

Parameter	Value
Network grid	Form (0,0) to (100,100)
Base station	At (200,50)
Threshold distance (d_0)	87 m
E_{elec}	50 nJ/bit/
E_{fs}	10 pJ/bit/m ²
E_{mp}	0.0013 pJ/bit/m ⁴
E_{DA}	5 nJ/bit/signal
Data packet size	4000 bytes
E_{TX}	50 nJ/bit/signal
E_{RX}	50 nJ/bit/signal
Initial energy	0.5 J/battery

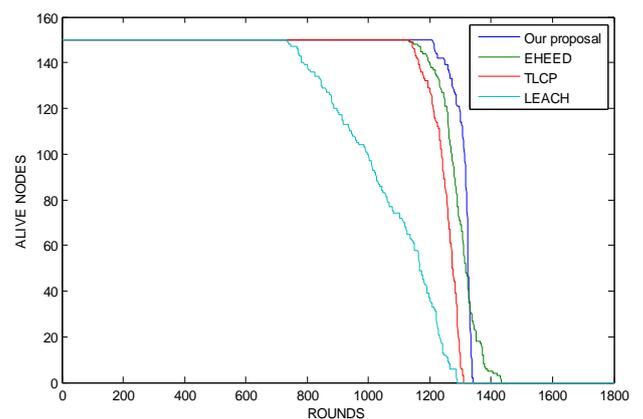


Fig. 6. The results for 1800 rounds running

As mentioned earlier, three metrics are used to qualify the network. FND (first node die) is very important so it is employed for application including on industrial or domestic gas and fire detection, this metric is used because dying one or more node might cause of mistake or danger on the case study.

Other metric parameter is HND (half node die) that demonstrate the number of round that half-number of all nodes die, until this round we still have a view of entire environment. LNA (last node alive) show the last round that there is a minimum view.

These parameters use to justify the correct operation of network. These parameters are simulated in MATLAB to compare the outcomes with other simulated algorithm results (Table 3).

TABLE 3. COMPARISON OF THE RESULTS

Algorithm	FND	HND	LNA
LEACH	735	1082	1291
TLCP	1141	1258	1311
EHEED	1133	1294	1433
Proposed algorithm	1210	1321	1340

Simulation result show that the proposal acts better than other algorithms in FND and HND, for LNA also is better than LEACH and TLCP (Fig. 6). According to graphs this is obvious that the average consumption of the proposed algorithm is less than the others.

4. CONCLUSION

In this paper, we have proposed an energy-efficient distributed clustering approach for sensor networks. In this approach, fuzzy logic is used to select the best nodes which they became cluster-head. The fuzzy system makes decision based on four parameters; residual energy as primary parameter, node proximity to its neighbors (centrality), distance to BS and node concentration. To minimize the network energy consumption, a multi hop method used for inter-cluster and intra cluster communication that causes the distance, which a message should send through it, is divided to several shorter paths. Comparative empirical study shows that the proposal extends network lifetime by selecting a suitable rate of CH and is less influenced with change the rate of CH. This algorithm can be tuned easily to optimize energy consumption according to selecting the best nodes for cluster heads.

5. REFERENCES

[1] Estrin D., Girod L., Pottie G., Srivastava M., 2001, Instrumenting the World with wireless sensor networks, International Conference on Acoustics, Speech, and Signal Processing , Salt Lake City, UT.
 [2] Butt I.M., and Khan S.A., 2005, Analyzing and Enhancing energy efficient communication protocol for wireless microsensor networks. ICICT 27–28, 323–327.

[3] Hedetniemi S., Liestman A., 1988, A survey of gossiping and broadcasting in communication networks. IEEE Networks,18(4), 319–49.
 [4] Heinzelman W., Chandrakasan A., and Balakrishnan H., 2000, Energy-efficient communication protocol for wireless microsensor networks. Proceedings of the 33rd Hawaii Int. Conf. on System Science (HICSS'00).
 [5] Afrashteh M.M., 2011, Design and Implementation a New Energy Efficient Clustering Algorithm using Genetic Algorithm for Wireless Sensor Networks. In World Academy of Science, Engineering and Technology 76, 430-433.
 [6] Gupta .I, Riordan G., and Sampalli S., 2005, Cluster Head Election Using Fuzzy Logic for Wireless Sensor Networks. In proceedings of IEEE Communication Networks and Services Research Conference, 255-260.
 [7] Younis O., Fahmy S., 2004, HEED: a Hybrid, Energy-Efficient, Distributed clustering approach For ad- hoc sensor networks. IEEE Transactions on MC2004, 3(4):366–79.
 [8] Sheikhpour R., and Jabbehdari S., 2012, A Two-Level Cluster based Routing Protocol for Wireless Sensor networks.International Journal of Advanced Science and Technology, Vol. 19-29.
 [9] Wang A., Yang D., Sun D. 2012, A clustering algorithm based on energy information and cluster heads expectation for wireless sensor networks. Elsevier, Computers and Electrical Engineering 38 , 662–671.
 [10] Heinzelman W., Chandrakasan A., Balakrishnan H., 2002 , An Application-Specific Protocol Architecture for wireless Micro sensor Networks. IEEE Transactions on Wireless Communications, vol. 1, no. 4, 660–670.
 [11] Banerjee S., Khuller S., 2001, A Clustering Scheme for Hierarchical Control in Multi-hop Wireless Networks. in Proceedings of IEEE INFOCOM,2001 .
 [12] Dargie W., Poellabauer CH., 2010, Fundamentals of Wireless Sensor Networks: Theory and Practice. Wiley Series on Wireless Communications and Mobile Computing , John Wiley & Sons, Ltd
 [13] Atslands R.R., 2012, WSNs clustering based on semantic neighborhood relationships. Elsevier Computer Networks 56 , 1627–1645.
 [14] Ivan S., Xu L., 2001, Power-aware localized routing in wireless networks. IEEE Transactions on Parallel and Distributed Systems,12(11),1122–33.