# Clustering Algorithm in Wireless Sensor Network: A Survey

Priti Kumari

***Abstract--In this survey paper, various challenges to wireless sensor networks are considered and Clustering as a solution to one of them has been introduced. The wireless sensor network is divided into clusters in order to reduce energy consumption as well as to increase coverage area to larger extent. The various clustering schemes with different parameters and in different environments such as static and mobile are discussed. The fuzzy approach in cluster head selection has overcome uncertainty of wireless sensor network at different levels.***

***Keywords-- wireless sensor network, clustering, cluster head, sensor node, hop, fuzzy approach***

**PritiKumari**
Computer Science Engineering
National Institute of Technology
Patna, India
priti_kaushik2003@yahoo.com

## I. INTRODUCTION

The easy deployment in hazardous environment, miniaturization of sensor devices, and wireless connectivity with base station and other similar nodes and mobility in order to increase coverage area have made wireless sensor network (WSN) a prominent choice in various application domains. Environment monitoring, battlefield surveillance, smart home, health analysis, agriculture are some examples [1]. With the use of sensor nodes to measure temperature, pressure, humidity, light conditions etc early precautions against draught, frost, and flood in agriculture might be taken. Sensor networks face a challenge of irregular topology as they are deployed randomly in the area of attention where set-up and maintenance of network is done autonomously. For the organization of monitoring of the target the whole area is divided into small units called clusters as shown in Fig 1. Clustering is an efficient technique for reducing energy consumption in the designed network [2] in which a leader Cluster Head (CH) is selected to control the activity within the cluster or to communicate with other CHs or base station. In the further discussion, the proposed algorithms for Cluster Head selection will be highlighted. Here, our objective is to use clustering for reducing energy consumption leaving other benefits as secondary. There is no facility of recharging or replacing of the batteries of nodes once they are deployed; therefore energy scarcity becomes a major issue during sensing, processing and communication. In such an environment, energy expenditure should be minimized at three levels: microprocessor-level, node-level and network-level [3]. All parts of a sensor, in addition to the network as a whole, must work together to lessen energy consumption in order to form an energy optimized sensor network.

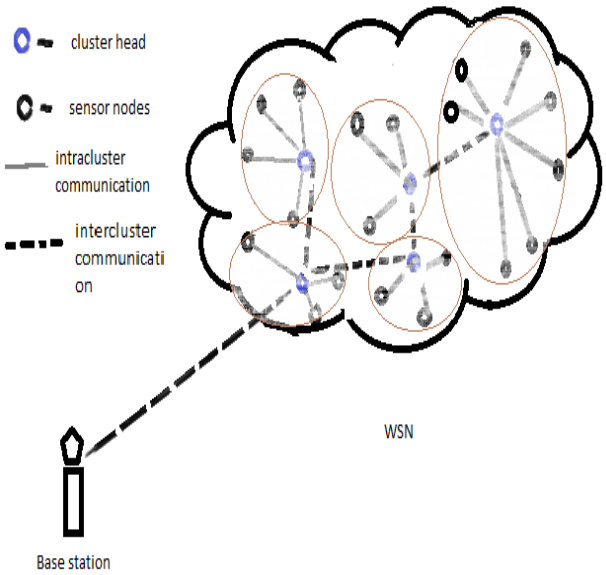


Figure 1. Clustered wireless sensor network

*WSN challenges—*

Since wireless sensor nodes have no facility of recharging their cells as they are deployed in diverse situations, energy conservation becomes main issue in order to keep alive the sensor network. Suppose a node dies in a particular area, that area might remain unmonitored unless that node is replaced by another energized node. Hence for efficient working of network and for proper outcome from the sensed data the energy need to be consumed in each process at each level. For example, [1] at MAC layer energy got wasted in different forms like packet collisions, idle listening, control packet overhead. The nodes especially those come on forwarding paths face serious battery depletion and proper protocols for collision avoidance requires enough energy. The above example is related to only one part of wireless sensor network but at each part such as routing, clustering, sensing, processing energy consumption is needed.

Sensor networks need to be self-configurable and scalable as there is need to replace damaged or discharged nodes form working nodes. And in the case of mobile sensor networks, mostly targets need to be monitored with scaling architecture because it is difficult to put sensor nodes on the particular target location. The nodes need to adjust themselves in various different conditions aroused according to environment. The architecture of network should be designed by sensor nodes in such a way that single node failure should not affect the working of whole network. It becomes tedious task with infrastructure less architecture and a limited energy resource as system is constrained to communicate with limited messages in order to set-up the topology among them.

Wireless sensor networks possess the challenge of security as they have liability of keeping data and themselves safe from external attacks. Although they use encryption scheme for data safety, though the full security of data means prevention from unauthorized access and alteration, maintaining data integrity requires security scheme to be applied at broad level which consumes a high energy. The probability of denial of service is also there as the communication line is kept busy, thereby true information cannot be sent to the desired node resulting in drainage of energy.

While taking a look over several challenges of wireless sensor network, energy management has been considered as most prominent concern.And Clustering is taken here as a solution to most of those challenges.

## II. TERMINOLOGY

1. *Wireless sensor network:* **A** wireless sensor network is a group of sensor nodes and base station which is designed to sense the environmental conditions like temperature, pressure, humidity, sound, light, air speed, direction, movement, vibration etc.. A wireless sensor network may be bi-directional implying that sensors send data to base station or main station and base station might control the activity of sensors.
2. *Sensor Node:* A sensor node is the core component of a WSN. Sensor nodes can take on multiple roles in a network, such as simple sensing; data storage; routing; and data processing.
3. *Base Station:* The base station is at the upper level of the hierarchical WSN. It provides the communication link between the sensor network and the end-user. The base station is assumed to have unlimited power, processing capability and storage capacity.
4. *Clustering*-Clusters are the organizational unit for WSNs. The dense nature of these networks requires the need for them to be

broken down into clusters to simplify tasks such a communication.

5. *Cluster head*- Cluster heads are the organization leader of a cluster. They often are required to organize activities in the cluster. These tasks include but are not limited to data-aggregation and organization the communication schedule of a cluster. A network organizer can choose a node as Cluster Head or there are algorithms form which sensor nodes can elect their cluster head.
6. *Hop*-Hop is total number of sensor nodes between source node and Cluster Head including source sensor node. If source node directly communicates with Cluster Head it is referred as single hop and if there are intermediate nodes, it is multi hop.
7. *Node degree*- node degree is the number of nodes with which a particular node is connected.

# III. TYPES OF CLUSTERING

## A. *BASED on authority-*

- *Centralized clustering scheme*- In the centralized clustering scheme base station keeps the right of determining the Cluster Head by judging the battery level and node location [4].

*Fuzzy-logic for Cluster Head selection*

In [5], the authors have taken three parameters for defining fuzzy rules: node concentration, energy level at each node and node's centrality with respect to entire cluster. As this algorithm comes under centralized algorithm, the cluster head is chosen by base station on the basis of earlier defined parameters. Similar to LEACH this scheme also contains two steps i.e. set-up phase and steady phase. In set-up phase base station selects cluster head and in steady state phase the cluster head processes the data sent by other nodes and after compressing it sends to base station.

In set up phase the fuzzy inference engine uses Mamdani method with three input variables energy, concentration and centrality. For node concentration and energy three linguistic variables are taken low, medium and high and to calculate centrality base station calculates the square of distance between selected node and other nodes in cluster because transmission energy is proportional to square of distance. There are three levels for node centrality close, far and adequate and seven types of chances for being elected as Cluster Head i.e. very small, small, rather small, medium, rather large, large, and very large. With the fuzzy rules, three input variables, the outcome chance of being a node as Cluster Head comes. The node with the highest probability is selected as Cluster Head.

- *Distributed clustering scheme* – In distributed clustering scheme sensor nodes take autonomous decision of choosing Cluster Head without intervention of base station.

*HEED (Hybrid energy-efficient distributed)* - HEED (Hybrid Energy-efficient distributed) clustering scheme is a clustering scheme for homogenous network where hybrid refers to the balance between energy and communication cost [6]. HEED uses residual energy as primary parameter because our main concern is to prolong network lifetime leaving node degree and distance to neighbors as secondary parameters.

Assumethat there are 'N' nodes in a wireless homogenous sensor network. With the primary clustering parameter i.e. residual energy, we select a set of possible cluster heads and with secondary parameter we choose one Cluster head from the set i.e. the cluster head with lower intra-cluster communication cost are favored. As energy required for sensing, processing and communicating one bit information is known thereby we can estimate the residual energy. At the end of HEED protocol either a node is chosen as Cluster Head or remains a regular node.

## B. *Based on functionality and characteristic of sensor nodes-*

With the design and feature of wireless sensor networks clustering algorithm also varies. There are

two types of networks: homogenous and heterogeneous network.

- In homogenous wireless sensor networks, all the sensor nodes have identical capacity, computing power, sensing range and hardware complexity. In such a network the CH dies before other nodes as it possesses the extra responsibility of transmitting information up to base station as well as data aggregation and processing.

*LEACH (Low Energy Adaptive Clustering Hierarchy)*

LEACH is one of the most accepted distributed cluster-based routing protocols in wireless sensor networks. LEACH allows the use of several data aggregation techniques within the cluster in order to get more précised data for forwarding, as it relies on the concept that nodes closer to each other contains correlated data [7]. LEACH selects a few nodes as cluster heads and rotates this role to balance the energy dissipation of the sensor nodes in the networks. As we know in the network, Cluster Head becomes over-burned and prone to be drained fast than others, therefore LEACH operates in rounds to rotate the cluster head responsibility [8].

Suppose there are N numbers of nodes in the network and k is the expected number of cluster heads in a round. Each node needs to be cluster head in N/k rounds with the purpose of maintaining energy balance in the network, therefore there is an indicator function Ci (t)to keep count that was node 'i' recently Cluster Head or not i.e. Ci (t)=0 if node was recently a cluster head otherwise 1 . Hence probability Pi (t) of being chosen as a Cluster head of node 'i' at round r is given as below:

$$P_i(t) = \begin{cases} \dfrac{k}{N - k*\left(r \bmod \frac{N}{k}\right)} & : \quad C_i(t) = 1 \\ 0 & : \quad C_i(t) = 0 \end{cases},$$

……..Equation (1)

The node with highest probability is chosen as Cluster Head. Once a node is chosen as Cluster Head, it broadcasts advertisement message to all other nodes in the cluster using CSMA protocol. The other nodes decide that in which cluster they want to reside on the basis of received signal strength from cluster heads.

*Fuzzy-logic based clustering approach-*

This clustering scheme is based on energy prediction, unlike other clustering scheme it considers expected residual energy as a fuzzy descriptor where other clustering schemes were considering residual energy and rotation of cluster heads. The clustering includes two phases: set-up phase and steady-state phase as shown below.

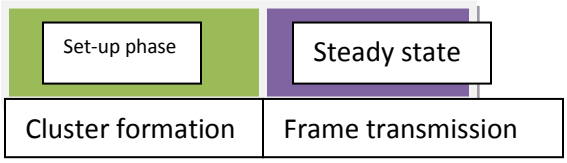


Figure2. Phases of clustering

This algorithm can also be called as extension of LEACH with expected energy residual where inference engine takes into account two parameters such as residual energy and expected residual energy to decide the probability of being chosen as cluster head. Suppose at time't' a node has RE (t) residual energy, then we can calculate ERE (expected residual energy) after time 'r' with the following formula:

$$ERE\ (t+r) = RE\ (t) - EEC\ (r) \quad \ldots\ldots (2)$$

Where EEC (r) is Expected Energy Consumption during time 'r' which can be computed from earlier known energy consumption of sensing nodes.

- In heterogeneous network, there are different types of node in respect to their functionality and hardware complexity. Some nodes are specially designed to perform as a Cluster Head (CH) with more battery power and more computing and communicating ability. In that network cost of designing the network is less than homogenous

network because only few nodes are established or deployed with hardware capabilities of being Cluster Head whereas rests are general nodes of low cost. But the deployment of nodes needs a special care to cover the whole area as only some nodes can act as a Cluster Head [9].

*EEHC (Energy efficient heterogeneous clustered scheme)* - EEHC is designed to choose a Cluster Head in uniformly distributed heterogeneous wireless sensor network. This algorithm uses the concept of optimal number of clusters formation with the calculation of signal to noise ratio and energy dissipation with distance because it has been seen that if less number of clusters are formed than optimal size, then energy consumption increases exponentially [10]. For the selection of Cluster head, firstly initial energy of node is taken into account [11]. As heterogeneous network is being dealt, suppose there are three types of nodes out of 'n' nodes: super node having b times more power than normal nodes, advance node having 'a' times more power than normal nodes, normal node. As there are three different types of nodes and with different powers, the node will be alive more than general clustering network. For example EEHC scheme increases the lifetime of network by 10% in comparison to LEACH [12].

## C. *Based on communication between Cluster head and sensor nodes*

- *Single hop clustering scheme*

- In a single hop network sensor nodes use only one hop to contact the cluster head. The nodes which are at large distance from cluster head need to consume more energy for transmitting data to Cluster Head with respect to nodes which are closer to the Cluster Head. LEACH can be used for single hop network.

*Fuzzy clustering approach for mobile sensor nodes—*

In [13], dynamic clustering is used to deal with mobile sensor nodes. The algorithm mainly emphasizes on mobility management and decision making of a node for choosing its best parent. Beside the Cluster head (CH) there is a concept of Zone Head (ZH) where Zone is a collection of one or more clusters. If a ZH receives a join message from any node (n), it includes it in neighbors and checks that if it can provide its services to node n. if the ZH is able to provide services it joins node n otherwise ask anyone of its node having less reliability to the connection with ZH than node n to join other ZH. If other node becomes ready to join other ZH, the node n is added to set of cluster heads. The same clustering process is followed in case of CHs. At the node level, node n will send hello_message to node c and add node c to its neighbors, after that node n will send join_message to node c. if the node c is CH and able to keep node n it will send ok message, otherwise join_other message will be sent.

During advertisement of a ZH or CH, a node may receive more than one request to join cluster or zone. In such a situation, a node uses fuzzy logic to decide best parent node. There are four fuzzy logic descriptors: Energy level of the node (Battery charge), Mobility, Quality of Link - QoL(Reliability between a node and his parent) and the Quality of Received Signal – QoRS(Received Strength Signal Indicator) with three possible values low, medium, high.

- *Multi hop communication*

Ina multi –hop wireless network, there may be intermediate nodes between sensor nodes and cluster head. It provides an advantage towards scalability of network. It surely decreases the energy consumption but simultaneously creates delay in messaging. In case of multi hop sensor network, the cluster head near to base station bears extra burden of transforming information from other cluster head to base station which is called unequal clustering [13]. To overcome such problem, we form clusters of small size near to base station in comparison to clusters which are far from base station.

*Multi-hop LEACH*

Multi-hop LEACH is used in multi hop sensor network when the network is extended beyond a certain level. It is an extension of LEACH protocol. In set-up phase the cluster head is selected like

LEACH but in steady state phase the Cluster Head far from base station transfers its information to other Cluster Head [14].

## iv. Conclusion

Clustering provides an efficient solution of reducing energy consumption up to some extent, but there is very little number of algorithms in the area of mobile sensor network. Although, there is a proper algorithm for mobile sensor networks discussed above which provides a suitable Cluster head for a node, though energy utilization can be minimized by replacing two parameters Quality of link, Quality of received signal with direction of mobility. If the direction of a Cluster Head is away from centre, the cluster head need to be replaced as quality of link and quality of received signal with other nodes will decrease obviously. In dynamic clustering, energy consumption for clustering is increased due to regular change of topology of sensor nodes.

## *Acknowledgement*

The author would like to acknowledge Professor Dr M P Singh for his guidance and valuable suggestions in the area of Mobile Wireless Sensor Networks.

## *References*

[1] Daniele Puccinelli and Martin Haenggi, Wireless Sensor Networks: Applications and Challenges of Ubiquitous Sensing, Third Quarter 2005 IEEE Circuits and Systems Magazine

[2] Heinzelman, W., Chandrakasan, A., Balakrishnan, H.: An application specific protocol architecture for wireless micro sensor networks', IEEE Trans. Wirel. Commun., 2002, 1, 4), pp. 660–670

[3] http://www.libelium.com/development/articles/102632630827

[4] NoritakaShigei, Hiromi Miyajima, Hiroki Morishita, Michiharu Maeda, Centralized and Distributed Clustering Methods for Energy Efficient Wireless Sensor Networks, Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, March 18 - 20, 2009, Hong Kong

[5] I. Gupta, D. Riordan and S. Sampalli: Cluster-head Election using Fuzzy Logic for Wireless Sensor Networks. Faculty of Computer Science, Dalhousie University

[6] O. Younis and S. Fahmy, "Distributed Clustering in Ad-hoc Sensor Networks: A Hybrid, Energy-Efficient Approach," in *Proceedingsof IEEE INFOCOM*, March 2004, an extended version appears in IEEE Transactions on Mobile Computing, vol. 3, issue 4, Oct-Dec2004.

[7] http://www2.aegean.gr/dgavalas/en/iframe_files/papers/2009/Clustering-chapter.pdf

[8] D. Hall, *Mathematical Techniques in Multisensor Data Fusion*.Boston, MA: Artech House, 1992.

[9] W. Heinzelman, "Application-specific protocol architectures for wireless networks," Ph.D. dissertstion, Mass. Inst. Technol., Cambridge, 2000.

[10] VivekKatiyar, Narottam Chand, SurenderSoni, Clustering Algorithms for Heterogeneous Wireless Sensor Network: A Survey, International Journal of Applied Engineering Research, DINDIGUL Volume 1, No 2, 2010

[11] W. Heinzelman, A. Chandrakasan,H. Balakrishnan, Energy-efficient communication protocol for wireless microsensor networks, in: Proceedings of the 33rd Hawaii International Conference on System Sciences (HICSS'00), January 2000.

[12] Dilip Kumar, Trilok C. Aseri, R.B. Patel, EEHC: Energy efficient heterogeneous clustered scheme or wireless sensor networks, Computer Communications, 2008 Elsevier

[13] S. Bandyopadhyay, E. Coyle, An energy efficient hierarchical clustering algorithm for wireless sensor networks, in: Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003), San Francisco, California, April 2003.

[14] ShahramNourizadeh, Y.Q. Song, J.P. Thomesse, A Location-Unaware Distributed Clustering Algorithm for Mobile Wireless Networks Using Fuzzy Logic,7th IFAC International Conference on Fieldbuses & Networks in Industrial & Embedded Systems - FeT'2007, Toulouse : France (2007)