CLUSTERING IN WSNs: A REVIEW

Chavan Pragnesh H
Department of Electronics and Communication
Hasmukh Goswami College of Engineering, Ahmedabad, India

Abstract: In order to collect information more efficiently, wireless sensor networks (WSNs) are partitioned into clusters. Clustering provides an effective way to prolong the lifetime of WSNs. Current clustering approaches often use two methods: selecting cluster heads with more residual energy, and rotating cluster heads periodically, to distribute the energy consumption among nodes in each cluster and extend the network lifetime. However, most of the previous algorithms have not considered the expected residual energy, which is the predicated remaining energy for being selected as a cluster head and running a round. A fuzzy-logic-based clustering approach which will increase the lifetime of WSNs.
Keywords: clusterings, Wireless Sensor Networks ,fuzzylogic, optimization, lifetime of network.

I. INTRODUCTION

A. Wireless Sensor Network

Wireless sensor network is a network which can be made up of autonomous sensors and these sensors are spatially distributed where they are used to sense environmental conditions within the network. These sensed data are passed to the main location through the network. Such networks are bidirectional and enables sensor activity control. Nowadays the WSNs are used in many applications like process management, healthcare monitoring, industrial monitoring, environmental and earth sensing, combat field surveillance and so on. The WSN consists of several sensor nodes, where each sensor node is connected with another. Each sensor node includes several parts as in Fig

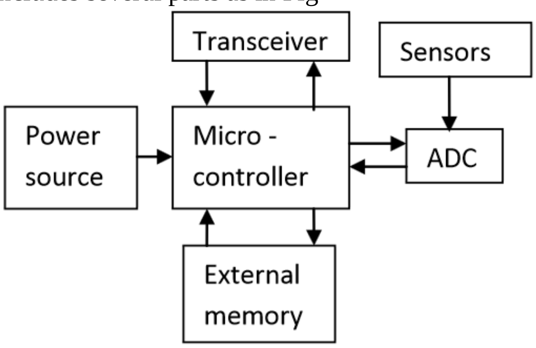

Figure **Error! No text of specified style in document.** Sensor node architecture

Sensors sense the data from the environment and those sensed data are converted into digital format by using analogue to digital converter. A microcontroller is an electronic circuit, acts as an interface between power source and sensors. A transceiver is used to transmit and receive the data to and from the other nodes in the network and an external memory which is used to store the data. Mostly, sensors are used to monitor sensitive data like enemy movement on the protected area. Therefore security is the major challenging task in wireless sensor network applications. They are low computation capability, limited memory, limited energy resources, use of insecure channel, sensitivity to physical capture, frequent changes in topology of network, sensor nodes are thickly deployed, sensor nodes are liable to failures, sensor nodes are not having id globally and number of sensor nodes can be varied in size.

B. Constraints in WSNs

Energy constraints: Energy consumption for the sensor nodes transducer, communication with other nodes and computation in microcontroller. Every bit transmitted in the WSN consumes energy more than the energy required for executing thousand instructions. Since the energy consumption is higher for the communication than the computation. Computation constraints: Processors in the sensor nodes have low computation capability than wired networks. Therefore complex cryptographic algorithms not used for the computation. Memory constraints: Sensor node's memory has RAM and flash memory. Purpose of RAM is to store sensor data, application programs and computations. Purpose of flash memory is to store downloaded application code. After the OS and application code execution, there is no space for executing complex cryptographic algorithms. Communication constraints: Range of communication is based on the receiving signal strength. Signal strength depends on the environment.

II. CLUSTERING IN WSNs

In direct communication WSN, the sensor nodes directly transmit their sensing data to the Base Station (BS) without any coordination between the two. However, in Cluster-based WSNs, the network is divided into clusters. Each sensor node exchanges its information only with its cluster head (CH), which transmits the aggregated information to the BS. Aggregation and fusion of sensor node data at the CHs cause a significant reduction in the amount of data sent to the BS and so results in saving both energy and bandwidth resources. Once the clusters are constructed, each sensor node will be given an exclusive time slot; therefore, each sensor node knows when to transmit. Consequently, a node does not require being awake. It is particularly crucial for scaling the network to hundreds or thousands of nodes. In many applications, cluster organization is a natural way to group spatially close sensor nodes in order to exploit the correlation and eliminate the redundancy that often shows up in the sensor readings.

Objectives for network clustering

Load balancing criteria
Objective for the case where CHs do data processing and duties of intra-cluster management are that distribution of sensor nodes among clusters evenly since it reduces data delay. To achieve the expected performance goals CHs has to balance the load among clusters and to extend the lifetime of the network have to set an equal-sized clusters since it prevents energy consumption of a subset of a cluster heads.

Fault tolerance techniques
The WSNs usually operated at the harsh environments and thus the nodes can easily get damaged. Therefore it is necessary to tolerate the failure of CHs for avoiding the loss of significant data in such applications. Re -clustering the network is one of the solutions to recover the failure of CHs, but this solution is not suitable for the on- going operation. For this reason use contemporary fault - tolerance technique that is assigning backup CHs to recover. Whenever the variations occur in the normal network operation the backup CHs have to play their role. The leaf nodes join the cluster based on the receiving signal strength and when CHs does not include nodes because of the long radio range then the neighbouring CHs can adapt those disjointed nodes. Solve the issues in fault-tolerance and load balancing in a cluster by means of rotating the role of CHs among sensor nodes in that cluster.

Maximize connectivity and minimize delay:
Inter-CH connectivity plays an important role while selecting the CH among sensor nodes in a cluster. The goal of connectivity is to ensure the availability of path between CHs to base-station. When assigning the CH role to some sensors, the objective of connectivity makes one of many variations in K-dominating set problem that is K-hop clustering. Intra-cluster connectivity is responsible for data latency. Factoring the delay usually based on maximum number of hops K in a path.

CH deployment
Specifically, the resource-rich nodes can be selected as CHs such as laptop computers, robots and mobile vehicle. The designer wanted to employ the smaller number of nodes since their size, cost, vulnerability and complexity of deployment.

III. WORK IN CLUSTERING IN WIRELESS SENSOR NETWORKS

In wireless sensor networks, the power resource of each sensor node is limited. Minimizing energy dissipation and maximizing network lifetime are important issue in the design of routing protocols for sensor networks.

A. Low Energy Adaptive Clustering Hierarchy in Wireless Sensor Network (LEACH)
Denso LEACH protocol is the first protocol of hierarchical routing which proposed data fusion; it is of milestone significance in clustering routing protocol. LEACH (Low Energy Adaptive Clustering Hierarchy) is designed for sensor networks where an end-user wants to remotely monitor the environment. In such a situation, the data from the individual nodes must be sent to a central base station, often located far from the sensor network, through which the end-user can access the data. There are several desira Conventional network protocols, such as direct transmission, minimum transmission energy, multi-hop routing, and clustering all have drawbacks that don't allow them to achieve all the desirable properties. LEACH includes distributed cluster formation, local processing to reduce global communication, and randomized rotation of the cluster-heads. Together, these features allow LEACH to achieve the desired properties. Initial simulations show that LEACH is an energy-efficient protocol that extends system lifetime. LEACH protocol is a typically representation of hierarchical routing protocol. It is a self adaptive and self organized. Leach protocol uses round as unit, each round is made up of cluster set-up stag e and steady state storage for the purpose of reducing unnecessary energy costs. The steady-state phase duration is usually much longer than set-up phase duration. However, the first phase is more important, in which sensor nodes are allowed to elect themselves as cluster-heads randomly, and then divided into clusters. Each node that becomes the cluster head (CH) will create a TDMA schedule for the sensor nodes within the cluster. That allows the radio components of each non- CH-node to be turned off all times except during their transmit time. Fig. 2 shows the cluster formation algorithm of LEACH.

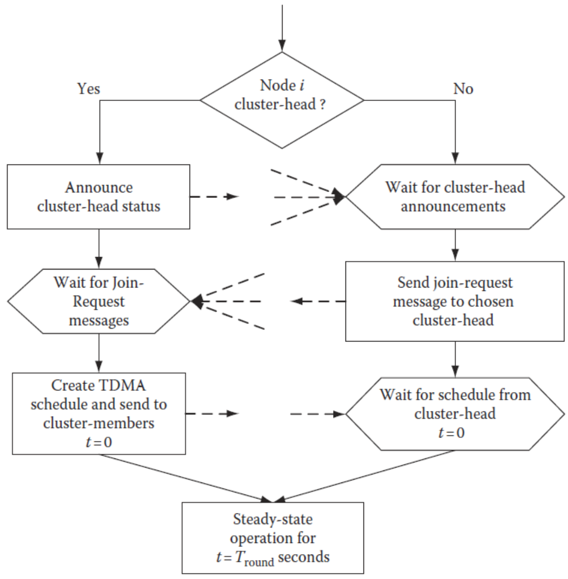

Figure 2 Flowchart of LEACH

B. Energy-Efficient Hierarchical Clustering (HEED)
For wireless sensor networks with a large number of energy-constrained sensors, it is very important to design a fast algorithm to organize sensors in clusters to minimize the energy used to communicate information from all nodes to the processing center. In this paper, we propose a fast, randomized, distributed algorithm for organizing the sensors in a wireless sensor network in a hierarchy of clusters with an objective of minimizing the energy spent in communicating the information to the information processing center. HEED (Hybrid Energy-Efficient Distributed clustering), has four primary goals:

- Prolonging network life-time by distributing energy consumption.

- Terminating the clustering process within a constant number of iterations/steps.
- Minimizing control overhead (to be linear in the number of nodes).
- Producing well-distributed cluster heads and compact clusters. HEED does not make any assumptions about the distribution or density of nodes, or about node capabilities, e.g., location-awareness.

In classical distributed systems, a node can either be a server or a source, but not both. A fixed number of servers is known to every source in the system, and a server is always available for processing. In our model, every node can act as both a source and a server (cluster head), which motivates the need for efficient algorithms to select servers according to the outlined system goals. A node only knows about the servers that are within its reachable range, which implies that achieving global goals cannot always be guaranteed but can be achieved through intelligent local decisions. Finally, a node may fail if its energy resource is exhausted, which motivates the need for rotating the server role among all nodes for load-balancing. HEED algorithm periodically selects CHs according to a combination of two clustering parameters. The primary parameter is their residual energy of each sensor node (used in calculating probability of becoming a CH) and the secondary parameter is the intra-cluster communication cost as a function of cluster density or node degree (i.e. number of neighbours). The primary parameter is used to probabilistically select an initial set of CHs while the secondary parameter is used for breaking ties. In HEED, the clustering process at each sensor node requires several rounds. Every round is long enough to receive messages from any neighbour within the cluster range. As in LEACH, an initial percentage of CHs in the network Cprob, is predefined. The parameter Cprob is only used to limit the initial CH announcements and has no direct impact on the final cluster structure. In HEED, each sensor node sets the probability CHprob of becoming a CH as follows

$$CH_{prob} = C_{prob} . \frac{E_{residual}}{E_{\max}}$$

where E residual is the estimated current residual energy in this sensor node and Emax is the maximum energy corresponding to a fully charged battery, which is typically identical for homogeneous sensor nodes. The CHprob value must be greater than a minimum threshold pmin. A CH is either a tentative CH, if its CHprob is <1, or a final CH, if its CHprob has reached 1.During each round of HEED, every sensor node that never heard from a CH elects itself to become a CH with probability CHprob. The newly selected CHs are added to the current set of CHs. If a sensor node is selected to become a CH, it broadcasts an announcement message as a tentative CH or a final CH. A sensor node hearing the CH list selects the CH with the lowest cost from this set of CHs. Every node then doubles its CHprob and goes to the next step. If a node completes the HEED execution without electing itself to become a CH or joining a cluster, it announces itself as a final CH. A tentative CH node can become a regular node at a later iteration if it hears from a lower cost CH. Here, a node can be selected as a CH at consecutive clustering intervals if it has higher residual energy with lower cost. In HEED, the distribution of energy consumption extends the lifetime of all the nodes in the network, thus sustaining stability of the neighbor set. Nodes also automatically update their neighbor sets in multi-hop networks by periodically sending and receiving messages. The HEED clustering improves network lifetime over LEACH clustering because LEACH randomly selects CHs (and hence cluster size)., which may result in faster death of some nodes. The final CHs selected in HEED are well distributed across the network and the communication cost is minimized. However, the cluster selection deals with only a subset of parameters, which can possibly impose constraints on the system. These methods are suitable for prolonging the network lifetime rather than for the entire needs of WSN.

C. The Gupta Fuzzy Protocol

A number of researchers have used Fuzzy Logic to extend network lifetime and minimize the energy consumption of the network. The Gupta protocol uses a Fuzzy Logic approach to select CHs. The FIS designer considered three descriptors: energy level, concentration, and centrality, each divided into three levels, and one output which is chance, divided into seven levels. The system also uses 27 IF-THEN rules. In this protocol there are two stages (set-up and steady-state) as in LEACH. The difference between the two protocols lies in the set-up stage where the BS needs to collect energy level and location information for each node, and evaluate them in the designed FIS to calculate the chance for each node to become a CH. The BS then chooses the node that has the maximum chance of becoming a CH.

After the CH selection, everything (advertising message, join CH message, and the steady-state stage) will be the same as in LEACH.

D. CHEF (Cluster Head Election mechanism using Fuzzy Logic)

The CHEF protocol (Cluster Head Election mechanism using Fuzzy Logic in Wireless Sensor Networks) [13], uses a Fuzzy Logic approach to maximize the lifetime of WSNs. It is similar to the Gupta protocol but it does not need the BS to collect information from all nodes. Instead the CHEF protocol uses a localized CH selection mechanism using Fuzzy Logic. Each node chooses a random number between 0 and 1. If this random number is smaller than Popt , the node calculates the chance using an FIS and advertises a candidate message with the chance. The message indicates that the node is a candidate for CH with the value of chance. Popt is calculated as:

$$P_{opt} = p \times \alpha$$

Where p is as in Equation (2.12) and α is a constant value that defines the ratio of the candidate for cluster head. The node then listens for candidate messages from nodes within radius *r* in equation,

$$r = \sqrt{\frac{area}{\pi.N.p}}$$

The node with the largest chance is selected as CH. After the CH selection, everything (advertising message, join CH message, and the steady-state stage) will be the same as in LEACH. The FIS uses two variables: energy residual and distance between nodes, one output and 9 IF-THEN rules.

E. LEACH-ERE (Expected Remaining Energy)

Before the cluster formation, the number of cluster members is unknown. However, since it is proportional to the number of neighbours near a potential CH (in a specific transmission range), the number of neighbours (defined as value *n)* could be used to obtain the expected energy consumption during the CH selection.

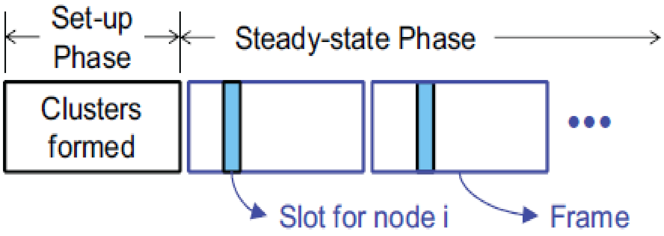

Figure **Error! No text of specified style in document.** Cluster formation and operation.

As shown in Fig., after the cluster formation, the steady-state operation is broken into frames, where nodes send their data to the CH at most once per frame during their allocated transmission slot. In a frame, suppose a CH has *n* cluster members, it would receive *n* messages from all the members and then transmit one combined message to the base station with a distance *d* to BS. The number of frames could be obtained by

$$N_{\text{frame}} = \frac{t_{\text{ssPhase}}}{n * t_{\text{slot}} + t_{\text{CHtoBS}}}$$

Where *t*ss Phase is the operation time of the steady-state phase (i.e. the time of a node to be a CH), *t* slot is the slotted time required for the transmission from members to the CH, and *t* CH to BS is the time required for the transmission from CH to the base station.

The expected consumed energy of a node to be a CH after a steady-state phase could be represented as

$$E_{\text{expConsumed}}(l, d_{\text{toBS}}, n) = N_{\text{frame}} * \big(E_{\text{Tx}}(l, d_{\text{toBS}}) + n * E_{\text{Rx}}(l)\big).$$

All the sensor nodes are assumed to transmit and receive the same size of messages, i.e. *l* bits of data. The distance to the base station, *d* to BS, could be computed based on the received signal strength. Then, the expected residual energy of a node to be a CH after a steady-state phase could be obtained via

$$E_{\text{expResidual}}(l, d_{\text{toBS}}, n) = E_{\text{residual}} - E_{\text{expConsumed}}$$

Where the *E* residual is the residual energy of a sensor node before the cluster head selection.

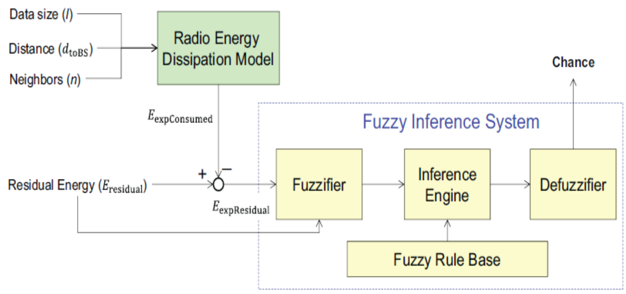

Figure 1 LEACH-ERE block diagram

IV. CONCLUSION

In this paper different clustering technique for wireless sensor networks is discussed. LEACH is the first protocol of hierarchical routing which proposed data fusion but LEACH does not ensure optimal lifetime of a network in terms of First Node Dies. After improvement in artificial intelligence algorithms researchers tried to implement them in WSNs clustering to fuse more than one parameter for cluster head election or for routing to enhance network lifetime. CHEF and LEACH-ERE has proved to provide better network lifetime than the LEACH and HEED.

REFERENCES

[1] Mhemed, Rogaia, et al. "An energy efficient fuzzy logic cluster formation protocol in wireless sensor networks." Procedia Computer Science 10 : 255-262, 2012.

[2] Kulkarni, R.V.; Forster, A.; Venayagamoorthy, G.K., "Computational Intelligence in Wireless Sensor Networks: A Survey," in Communications Surveys & Tutorials, IEEE , vol.13, no.1, pp.68-96, First Quarter 2011.

[3] Jin-Shyan Lee; Wei-Liang Cheng, "Fuzzy-Logic-Based Clustering Approach for Wireless Sensor Networks Using Energy Predication," in Sensors Journal, IEEE , vol.12, no.9, pp.2891-2897, Sept. 2012.

[4] AlShawi, I.S.; Lianshan Yan; Wei Pan; Bin Luo, "Lifetime Enhancement in Wireless Sensor Networks Using Fuzzy Approach and A-Star Algorithm," in Sensors Journal, IEEE , vol.12, no.10, pp.3010-3018, Oct. 2012.

[5] Handy, M. J., Marc Haase, and Dirk Timmermann. "Low energy adaptive clustering hierarchy with deterministic cluster-head selection." Mobile and Wireless Communications Network, 2002.

[6] Katiyar, Vivek, Narottam Chand, and Surender Soni. "Clustering algorithms for heterogeneous wireless sensor network: A survey." International Journal of Advanced Networking and Applications vol.2, no. 4, pp.745-754, 2011.

[7] Bandyopadhyay, Seema, and Edward J. Coyle. "An energy efficient hierarchical clustering algorithm for

wireless sensor networks." In INFOCOM 2003. Twenty-Second Annual Joint Conferences of the IEEE Computer and Communications. IEEE Societies, vol. 3, pp. 1713-1723, 2003.

[8] Kim, Jong-Myoung, Seon-Ho Park, Young-Ju Han, and Tai-Myoung Chung. "CHEF: cluster head election mechanism using fuzzy logic in wireless sensor networks." In Advanced communication technology, 2008. ICACT 2008. 10th international conference on, IEEE, vol. 1, pp. 654-659., 2008.

[9] Mhemed, Rogaia, Nauman Aslam, William Phillips, and Frank Comeau. "An energy efficient fuzzy logic cluster formation protocol in wireless sensor networks." Procedia Computer Science vol. 10, pp 255-262, 2012.

