Efficient Clustering techniques for WSN- a survey

Mamta Tewari[#1], Reenu Saini[*2]

Computer Science Department, Uttarakhand Technical University
BTKIT, Dwarahat, Almora, India
[1]mamta20tewari@gmail.com

Computer Science Department
BTKIT, Dwarahat, Almora, India
[2]reenurkesaini@gmail.com

***Abstract*— This paper gives an overview of some of the efficient clustering algorithm proposed for the wireless sensor network. As there are various potential applications of the WSN like in military surveillance, disaster management, smart homes, habitat monitoring where a huge number of sensor nodes are deployed in an unattended environment which requires the network lifetime to be prolonged. To extend the network lifetime, for high scalability a disjoint, non-overlapping subsets of sensor nodes is generally formed known as cluster. Clusters form hierarchical WSNs which aid in proficient utilization of resource constrained sensor network and hence increases network lifetime. The main aim of this paper is to give a state of art review on clustering algorithms proposed for WSNs.**

***Keywords*— sensor nodes; wireless sensor network (WSN); clustering; energy efficiency**

I. INTRODUCTION

The advances in Micro-electro-mechanical systems (MEMS) have formed a platform for the emergence of wireless sensor network (WSNs) and its wider usage in our lives. Smart sensors are power constrained devices that have one or more sensors, memory unit, processor, power supply and actuator [1]. Wireless Sensor Networks provide a wide range of applications from military applications such as battlefield monitoring and target surveillance, to creating smart homes where sensors can monitor safety and offer automated services to the individual user; there are endless applications of WSNs.
The sensor nodes have limited processing power, communication bandwidth and storage capacity which demands for proficient utilization of resources. In WSNs the nodes are regularly organized in a disjoint set known as cluster. To offer network scalability, sharing of resource, efficiency in using constrained resources clustering is used in WSNs. Clustering helps in reducing communication cost, total energy consumption of network , and also reduces the interferences among sensor nodes [2].
The figure 1 shows a general architecture for Wireless Sensor Network [1].

- Sensor Node: the basic component of a WSN is a sensor node. It does multiple work like sensing; storing data; processing data; and routing.
- Clusters: the organizational unit of a WSN is called as cluster.
- Clusterheads: the leader of a cluster is referred to as cluster head. They basically manage activities in the cluster.
- Base Station: It is the topmost level of a hierarchical WSN and offers communication between the end user and the sensor network.

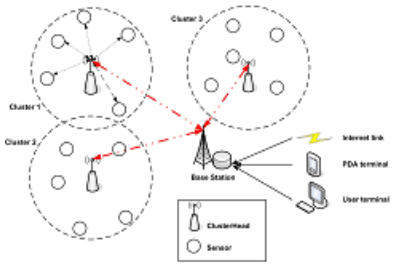

Fig 1 Architecture of sensor network

II. LITERATURE REVIEW

In recent years, the interest in clustered wireless sensor networks has produced a lot of research works. The cluster head (CH) of a cluster is elected by other sensor nodes dynamically or it can be pre assigned to a cluster initially at network set up stage by the network designer. The CH may be one of a node from the other cluster node or it may be a prosperous node in terms of resources.

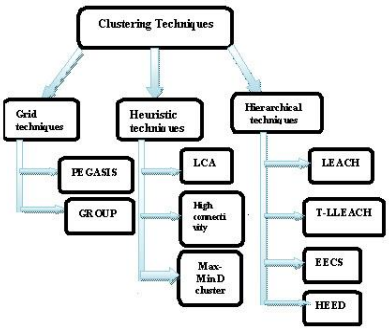

Fig 2 Taxonomy of proposed clustering algorithm

The diagram 2 gives the overview of all the clustering algorithms proposed so far for the WSNs [3]. We have focussed on three classification of clustering techniques i.e Heuristic, Grid and Hierarchical. We will briefly discuss the algorithms falls into the categories shown in the diagram and see the pros and cons of the various clustering techniques and analyse their performance.

A. GRID TECHNIQUES

1) *Power-Efficient GAthering in Sensor Information Systems (PEGASIS) [8]:* Basically a data-gathering algorithm based on chain protocol and gives improvement over LEACH protocol. Authors proposed the idea that a chain is formed from source node to sink node and only one node is transmitting to the base station node in any given time frame. This algorithm shows an improvement over LEACH [8] by bounding the total number of transmissions and thus improving energy consumption of algorithm over other clustering algorithms.

2) *GROUP [9]:* It is a grid-based clustering algorithm. In GROUP one of the sink node develops the cluster GRID [9]. As shown in fig 3 the nodes are arranged in the form of grid. The queries are forwarded from the base node to source node via the GridSeed (GS) to its cluster head of the node. The GS is a node within a given radius from the primary sink [9] and selection of GS is based on the residual energy by the primary sink.

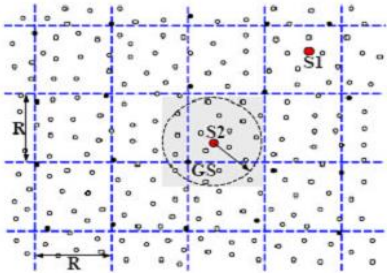

Fig 3 a cluster grid example of GROUP [10]

B. HEURISTIC TECHNIQUES

1) Linked Cluster Algorithm (LCA) [4][5][6] : It was the very first clustering algorithm proposed. Initially developed for wired networks, later on implemented for WSN. A unique ID number is given to each node and a node becomes a cluster head in two ways. In the first way if the node has the maximum ID number in the group containing all the neighbor nodes and itself. The second way is if the node assumes that none of its neighbor is a cluster head.

2) Highest Connectivity Cluster Algorithm [5]: in this scheme instead of seeing the ID number the connectivity of a node is considered. Cluster head is the node with highest connectivity, and in case if there is a tie then node with lowest ID wins.

3) Max-Min D-Cluster Algorithm [7]: As proposed in this algorithm the nodes are not more than d (a parameter taken) hops away from the cluster head node. Cluster head selection process is as follows: 2d rounds of flooding are started by each node from and results are considered from it. To select the cluster head node, each node follows some rules. The 1st d round called as fmax does the function of circulating the largest node IDs. After the completion of 1st d round, the 2nd d rounds of flooding follows known as fmin. A node declares itself cluster head if it receives its own ID after fmax round. The correctness of this algorithm depends on following two assumptions:

- No node ID will be circulated beyond d- hops from originating node during flooding.
- Nodes surviving the fmax choose themselves the cluster head.

C. HIERARCHICAL TECHNIQUES

1) LEACH: Low Energy Adaptive Clustering Hierarchical Protocol (LEACH) [10] was the first vital improvement over traditional clustering techniques. It provides the mechanism of random circulation of the clusterheads and hence helps in balancing the energy usage by the nodes [10]. To lessen the inter-cluster and intra cluster collisions LEACH uses the TDMA or a CDMA MAC. Various factors like the number and type of sensors, communication range and geographical location becomes the basis of cluster formation. The algorithm consists of two phases: set up phase and steady state phase. In set up phase the clusters are formed and information transfer takes place in steady state phase. The energy utilization of the data collected by the sensor node to arrive at the base will rely on the number of clusterheads and radio range of different algorithms, due to the reason that the energy utilization can be curtailed by arranging the sensor nodes in clusters [11].

2) Two-Level Hierarchy LEACH (TL-LEACH) [12]: It is an extended version of LEACH protocol and utilizes two levels of clusterheads. The algorithm consists of three phases: selecting two cluster-head, setup and steady state phase like in LEACH. Data transmission from source node

to sink is accomplished in two steps [12]: Secondary node gathers information from the nodes of their clusters and then transmits it to primary node.

The two level hierarchy decreases the number of nodes that transmits data to base node and hence potentially reduces the total energy consumption.

3) *Energy Efficient Clustering Scheme (or EECS) [13]:* In this algorithm the cluster build up phase is different from the LEACH. LEACH forms cluster on the basis of least distances of nodes from their respective cluster heads [14] but in EECS dynamic sizing of clusters depending on distance of cluster from the base occurs [13]. This algorithm focuses on the issue that the node which are far from the base station consume more energy than the node which are near.

The energy distribution is improved resulting in efficient use of resources with increasing lifetime of network.

4) *Hybrid Energy-Efficient Distributed Clustering (or HEED)* [15]: a multi-hop clustering algorithm in which the cluster head selection is based on the physical distance between nodes. The clusterhead selection is of prime importance in HEED and it is done based on two parameters: 1) The residual energy 2) Intra-Cluster Communication Cost.

The concept of residual energy is exploited in a number of clustering schemes. The Intra-Cluster Communication Cost is used by nodes to decide which cluster to adhere. This parameter is quite beneficial if a node falls in the territory of more than one clusterhead.

5) *Deterministic energy efficient clustering algorithm DEC [19]:* Authors in DEC[19], proposed a deterministic energy efficient algorithm in which the cluster head selection algorithm is solely based on the residual energy of the node [19]. They have proposed the algorithm for homogeneous as well as heterogeneous system. The algorithm works in two phases: set up phase and steady state phase. The steady state phase is similar to LEACH [10] but there is some change in the setup phase. The technique works in improving the stability region of the network with deterministically selecting the cluster head unlike the probabilistic based algorithm where the cluster head selection is probabilistic.

D. *Miscellaneous Clustering techniques:*

1) *Clustering Algorithm based on Cell Combination (CACC) [16]:* A clustering algorithm based on cell combination is proposed. The observed region is divided into hexagonal cells taking into account the geographic location of nodes. A cluster is formed with nodes having same cluster identity and at least seven hexagon cells are present in a cluster. The central cell of each cluster is considered to select the cluster head. In order to improve the energy efficiency and channel reuse the cells shape is considered circular.

III. PERFORMANCE ANALYSIS OF PROPOSED TECHNIQUES

The performance of a clustering algorithm is measured in terms of following parameters:

- Power
- Energy consumption
- Network lifetime
- Quality of network
- Reliability of links

As for heuristic algorithms, specific comparisons cannot be made in terms of power and quality, because these algorithms do not depend on some quantitative parameters.

A. In LCA [4], [5], [6] nodes use TDMA frames to communicate, that means LCA is feasible for small number of nodes only [18].

In Highest-Connectivity [5], a high change of clusterheads happens when the topology of network changes. So it is required that each node must store network topology information and this information is moved from previous cluster head to new cluster head when cluster head changes. It is a high overhead which limits the rotation of cluster heads.

In [7], d-hop Max-Min algorithm is proposed which create much larger cluster size with less number of cluster heads, and the duration of cluster head is longer as compared to LCA. A node in d-hop Max-Min remains as clusterhead approximately 100% longer than LCA2 and it goes on rising on increasing the density of network. [18]. Hence, the above quality of d-hop Max-Min shows that it is the best clusterhead election algorithm as compared other heuristic algorithms.

B. *PEGASIS [8]:* Energy minimization is obtained in this algorithms in the following ways:

In a given round a single node is sending data to the sink. The concept of communicating with the nearest neighbor helps in efficient utilization of energy. Data fusion is performed by each node, thus energy is distributed effectively across the network. Dynamic cluster formation overhead is also removed.

The reliability of PEGASIS is not as much promising as the network lifetime improvement is. It focuses on communication with nearest neighbor which can cause failure as there are gaps in a network.

C. *GROUP [9]:* In GROUP, energy is saved by following lower transmission distance for data sent to sink. In LEACH a clusterhead send information straight to base station, while in GROUP, data is send in short ranges instead of directly to sink.
The graph of simulation of LEACH vs GROUP is shown in figure 4. The energy consumption of LEACH and GROUP is comparable for node 75, but for greater value of nodes GROUP offers a reduction in the maximum energy consumption per node by a factor of 4 [17]. The energy distribution for a larger number of nodes is steadier with GROUP in comparison with LEACH.

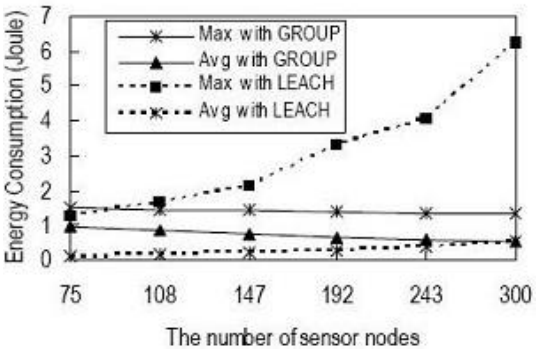

Fig 4 energy consumption curve for LEACH and GROUP [17]

In case if a node fails in communicating with the cluster head GROUP provides some recovery process. The node broadcast message to explore and launch a new cluster head. Failed node neighbors elect a new cluster head similarly to grid construction [17]. GROUP assures to transmit data from source to base station successfully.

D. LEACH and TL-LEACH [10][12]: The benefits of LEACH can be described in following points:
- No overhead in Cluster head selection process because each node decides in this process.
- CDMA aid in operation of node independent of each other.
- The minimum transmission energy is calculated by each node to communicate with its cluster head. So the transmission energy is reduced in LEACH
- The first node dies (FND) time is very late in LEACH as compared to other clustering techniques.

TL-LEACH [12]: It is an improvement over LEACH by implementing two level hierarchy. The energy of system is improved in TL-LEACH as in this few number of nodes transmit large distances. The network lifetime is also increased in TL-LEACH by 30% as compared to LEACH [12].
There are many features in LEACH and T-LEACH which ensure a reliable transmission.
- Collisions are avoided using CSMA technique.
- Neighboring node interferences in avoided using CDMA.
- The network lifetime is extended by rotating the role of cluster head from time to time.

E. EECS [13]: This algorithm also works on improvement over LEACH and energy consumption minimization is performed in the same way as in LEACH [13]. EECS improved the network lifetime by 35% on comparison with original LEACH.
EECS provides similar recovery mechanism as in LEACH. The reliability is achieved in EECS as the energy utilization is improved resulting in longer duration of network.

F. HEED [15]: The network lifetime is improved by minimizing the number of nodes competing for accessing the channel. The cluster heads are well distributed in HEED unlike in LEACH where there is random selection of cluster heads.

G. DEC [19]: The stability region is increased in DEC as compared to other probabilistic algorithm. A fixed number of cluster heads are selected in each round. The network lifetime is also improved in DEC.

IV. CONCLUSION

In this paper we have discussed some clustering techniques proposed for wireless sensor network and analyse their results on some performance parameters. We have taken three classification of sensor network and described some clustering techniques falls under these categories. The constrained energy of sensor nodes is main criteria in designing any technique for WSN [1]. Other factors, such as recovery and reliability, network lifetime are also considered.
Techniques presented in this paper shows improvement in terms of energy consumption, network lifetime, stability from the traditional algorithms for WSN. Authors in [12][14] have focused in energy improvement by efficiently selecting cluster head and in [9][15] a distribution of cluster head is created. In [19] authors proposed a deterministic approach for cluster head selection which selects a fixed number of cluster head

per round improve the network stability period. The reliability of sensor network is also acknowledged in different techniques by doing re-clustering happening at different time duration but it results in energy inefficiency and limits the lifetime of network.

REFERENCES

[1] Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," IEEE Commun. Mag., vol. 40, pp. 102-114,2002.

[2] M. Younis, M. Youssef and K. Arisha, "Energyaware management in cluster-based sensor networks", Computer Networks43 (5), 2003, 649-668.

[3] Vinay Kumar, Sanjeev Jain and Sudarshan Tiwari, "Energy Efficient Clustering Algorithms in Wireless Sensor Networks: A Survey"IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 2, September 2011

[4] D. J. Baker and A. Epheremides, "The Architectural Organization of a Moblie Radio Network via a Distributed Algorithm," IEEE Transactions on Communications, vol. Com-29, no. 11, November 1981.

[5] P. Tsigas, "Project on Moblie Ad Hoc Networking and Clustering for the Course EDA390 Computer Communcation and Distributed Systems," Manual for University Course.

[6] A. Amis, R. Prakash, T. Vuong, and D. Huynh, "Max-Min D-Cluster Formation in Wireless Ad Hoc Networks," IEEE INFOCOM, March 2000.

[7] A. D. Amish, R..prakash, T. vuong, and D. huynh "Max-min d-cluster formation in wireless ad hoc networks". *In Nineteenth IEEE Annual Joint proceeding on Computer and communication societies*, 1, 32–41.

[8] S. Lindsey,. And C.S. Raghavendra, C.S. "PEGASIS: Power-efficient gathering in sensor information systems". In IEEE proceeding on aerospace, 3, 1125-1130.

[9] Y. Liyang, M.W. Neng, Z. Wei, and Z. Chunlei "GROUP: A grid-clustering routing protocol for wireless sensor networks". In *IEEE International conference on Wireless Communications, Networking and Mobile Computing*, 1–5.

[10] Efficient Communication Protocol for Wireless Sensor Networks," Proceedings of the 33th Hawaii International Conference on System Sciences, 2000.

[11] H. Junping, J. Yuhui and D. Liang "A Time-based Cluster-Head Selection Algorithm for LEACH" IEEE Symposium on Computers and Communications, 2008, pp.1172-1178.

[12] V. Loscri, G. Morabito, and S. Marano, A Two-Level Hierarchy for Low-Energy Adaptive Clustering Hierarchy, DEIS Department, University of Calabria.

[13] Y. Mao, L. Chengfa, C. Guihai, and J. Wu "EECS: An energy efficient clustering scheme in wireless sensor networks". In *Proceedings IPCCC, IEEE 24th International*, 535–540.

[14] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Sensor Networks," Proceedings of the 33th Hawaii International Conference on System Sciences,2000.

[15] O. Younis and S. Fahmy, "HEED: A Hybrid Energy-Efficient Distributed Clustering Approach for Ad Hoc Sensor Networks," IEEE Transactions on Mobile Computing, vol. 3, no. 4, Oct-Dec 2004.

[16] L. Chang-RI, Z.. Yun, Z. Xian-ha, and Z. Zibo "A clustering algorithm based on cell combination for wireless sensor networks" In Second International Workshop on Education Technology and Computer Science, 2,74–77.

[17] L. Yu, N. Wang, W. Zhang, and C. Zheng, GROUP: A Grid-Clustering Routing Protocol for Wireless Sensor Networks, East China Normal University.

[18] A. Amis, R. Prakash, T. Vuong, and D. Huynh, "Max-Min D-Cluster Formation in Wireless Ad Hoc Networks," IEEE INFOCOM, March 2000.

[19] Aderohunmu F.A., Deng J.D., Purvis M.K., "A deterministic energy-efficient clustering protocol for wireless sensor networks", Proceedings of Seventh IEEE International Conference on Intelligent Sensors, Sensor Networks and Information Processing, pp.341 - 346, 2011

