Clustering in Wireless Sensor Networks- A Survey

[1]Khushboo Vijayvergiya, [2]Manoj Singh

[1]M.Tech Scholar, [2]HOD (Computer Science)
Department of Computer Science and Engineering,
Gurukul Institute of Technology, Kota Rajasthan, India

***Abstract*—Wireless Sensor Networks are a rapidly growing area, especially in terms of research and commercial purposes. The main concern when dealing with such networks is to transfer the coverage information from source to the sink with minimized energy consumption, prolonged network lifetime and maximum coverage. An effective step in achieving the mentioned requisites is clustering. Clustering aims to reduce the load of the nodes in a cluster by limiting their work only to sensing and passing of information to the cluster head that then performs tasks like filtering, transfer and fusion. Also a lot of more factors favor clustering in WSNs. This paper provides a brief survey discussing the some of the traditional clustering algorithms for WSNS and the recent developments in the field.**

***IndexTerms*—Wireless Sensor Networks, Clustering, Maximum Coverage, Minimized Energy Consumption, Prolonged Network Lifetime**

I. INTRODUCTION (HEADING 1)

Wireless Sensor Networks (WSN) consists of autonomous sensors distributed randomly or at fixed locations in the region of interest to cooperatively work on a specific task of monitoring the environmental or physical conditions. Coverage is the term used for measuring the capability of nodes to observe the given region of interest. The information collected from these nodes can then collectively be passed to a main station, also referred to as the sink. Figure 1 shows the working of a WSN. Initially motivated for the need of battlefield surveillance in military, WSNs are now widely used for applications like critical infrastructure surveillance, pollution monitoring, monitoring and control in industries, machine health etc.

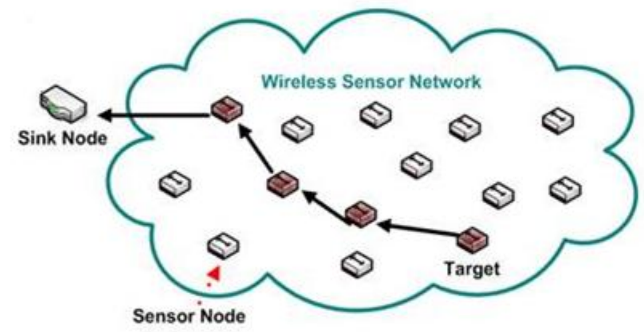
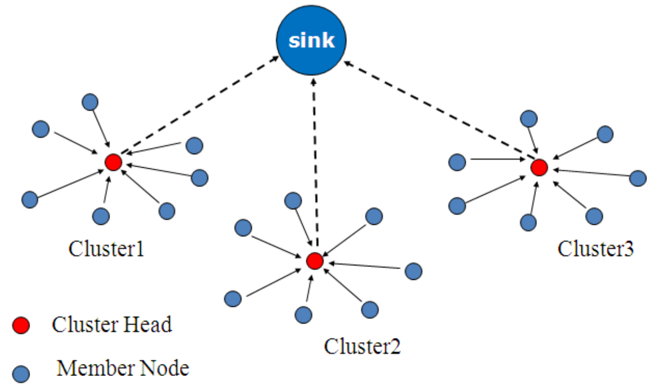


Fig 1Working of a WSN

Almost all research headed in the direction of WSNs assumes that the nodes in the network are of homogeneous nature, that is with same capabilities and energy resources. The other kinds of nodes are heterogeneous nodes with some nodes more powerful than the rest. Deployment of nodes can also be dense or sparse depending on the need of coverage and the resources available. The nodes participating in a WSN, whether heterogeneous or homogeneous, are dependent on a battery source for energy replacement or recharging of which is not feasible mostly and therefore mechanisms have to be set for achieving a minimized overall energy consumption of nodes along with maximized network lifetime. The other important requirement is to maximize coverage with a minimum number of nodes that can still ensure desired coverage. The criteria for coverage can be the nodes sensing and passing the information to the base station individually or sending the sensing information to the cluster heads which can then do the filtering, transportation and fusion tasks. In the case with nodes performing individually, each of the nodes' battery is consumed leading to an overall increased consumption rate. On the other hand, using cluster heads for tasks other than sensing, sufficient amount of energy of nodes can be conserved. The procedure for selecting cluster heads is done through clustering using clustering algorithms designed specifically for WSNs. This paper discusses some of the related traditional clustering algorithms followed by a brief survey of the developments in the area, both classic and recent.

II. CLUSTERING

Clustering in WSNs is a much researched area with hundreds of clustering algorithms available. Clustering, in data mining, is an unexploratory task aiming to group similar data clusters. In the clustering process in WSNs, each node can be organized in local clusters and one node of each cluster serves as a cluster head. The cluster head gets the sensing information collected by the non-cluster-head nodes aggregates it and sends it to the remote base station. The nodes in a cluster detect a similar phenomena and therefore data aggregation through cluster heads results in overall reduced information that is being sent to the base station. With only cluster heads involved now for the communication with the base station, the reduced communication bandwidth and reduced energy consumption with prolonged lifetime of individual nodes form two of the advantages of clustering in WSNs. The route setup in a cluster is now localized which reduces the size of the routing table. Otherwise each node has to update individually the routing table. The topology of the network can be stabilized at the level of nodes. The scalability of the network is increased and the involved nodes in a cluster do not interfere with the working of the other nodes. These all factors encourage clustering in WSNs for several tasks like energy efficient coverage, increased network lifetime etc. Through the years, there has been done significant work in this direction. Figure 2 shows the clustering process in WSNs. The member nodes communicate the cluster heads which then communicate with the base station.



III. Traditional Methods

The advancement towards clustering in WSNs is motivated by the previous research works done in this direction. One of the first and most recognized clustering algorithms for WSNs is the LEACH algorithm [1].

LEACH [1] algorithm by Chandrakasan et al is a hierarchical protocol with the transfer of information to the sink done by the cluster heads on the basis of the sensing information they receive from the other nodes in the clusters. Though assuming that all the nodes have a radio frequency powerful enough to help them reach the sink or the nearest cluster head, the consumption of the whole energy will still be a waste for the network. The nodes go through a stochastic algorithm to check if they have been assigned as cluster head at each round. Once the cluster has been a cluster head, the same cannot happen till P rounds where P represents the percentage desired for the cluster heads. A node can therefore have 1/P probability of becoming a cluster head. When the round comes to an end, the non-cluster-head-nodes become a part of the clusters with the closest cluster head. A schedule is then set by the cluster heads for each node in their respective cluster for data transmission purposes. The drawbacks of the algorithm are its strict time synchronization and no guarantee of a good cluster distribution. Also a lot of scalability issues arise when directly transmitting data from Cluster Heads to base station.

Lindsey and Raghavendra [2] then proposed an extension of the LEACH algorithm by the name PEGASIS. The methodology applied in PEGAGIS is that it organizes all the nodes in a chain and transmission of data is done only from a node to its nearest neighbor. This way the energy consumption of the nodes is significantly reduced. The effectiveness of the proposal is limited to small networks only with an observed high in the data delay for the large networks.

In the HEED algorithm [3], the parameter for cluster head election is the residual energy of nodes. When there is encountered a tie between any two nodes for cluster node election, then the algorithm checks the node degree or the average distance between the neighbors for election. The algorithm outperforms the LEACH algorithm in terms of its energy level consideration during cluster head election. A significant reduction in energy consumption with prolonged network lifetime as outputs proves the efficacy of the algorithm.

EECS [4] algorithm comes in one of the unequal clustering algorithms in WSNs. For the cluster head election, each node has to broadcast a control message to all the other nodes asking for competition. Every node has an associated radio range which if higher than the node sending the broadcast message enables the node to withdraw from the competition. Eventually the node with higher radio range is assigned the cluster head and it broadcasts another control message to announce the same. The algorithm also has an excellent criterion of dealing with energy consumption requirement of WSNs. The distant cluster heads have to consume energy for transmission of information to the base station. This problem is compensated by minimizing the number of nodes in the distant clusters. But in case of large scale sensor networks, early deaths of some nodes can occur due to a single hop transmission from the cluster heads to the base station leading to biased energy consumption between the cluster heads.

EEUC [5] algorithm also comes in the unequal clustering algorithms in WSNs like EECS. Both algorithms share a somewhat similar methodology, the only difference being the criterion for competition for the cluster head election. Whereas in the former algorithm, a higher radio range corresponds to the increased probability of the node of being assigned the cluster head; in EEUC, the distance of the node to the base station is evaluated. Higher distance enables a node to withdraw anyone form the competition and becomes the cluster head itself. A parameter TD.MAX is also tested which if higher than the distance between the cluster heads and the base station, leads to direct transmission. Otherwise, multi-hopping occurs without data aggregation.

IV. Latest Contributions

COCA[6], an unequal clustering algorithm proposed by Li et al addresses the issue of energy-hole problem in multi-hop clustering model with uniformly distributed nodes and equally sized clusters. Due to the heavier relay traffic, the cluster heads that fall closer to the sink consume more energy and may eventually lead to a 'hole' causing the network hole problem. The idea behind the proposal is to divide the entire region of interest into units. The nodes that have higher energy than their neighborhood nodes are elected as the cluster heads. Each unit can consist of more than one cluster. The size of the clusters away from the base station is larger as compared to the clusters closer to the base station. The performance of COCA is observed higher in cases of non-uniform node distributions.

The DSBCS[7] algorithm proposed by Liao et al consists of three phases to give as output balanced and unequal clusters. The first step of cluster head election lies solely on the weight of nodes examining which is further dependent on the residual energy of nodes and node connectivity. The second step of cluster building consists of the nodes joining the nearest cluster heads having high residual energy. The algorithm also limits the number of nodes that can become a part of the cluster. A threshold value is assigned crossing which no more node requests are entertained. The third and the last phase of the algorithm, the cycle phase, involves evaluation of weight of each of the member nodes in the cluster by the cluster heads for election of newer cluster heads. However, the authors retain the same cluster in the cycle phase avoiding the energy consumption occurred in the new cluster formation.

Gupta et al [8] extended the LEACH algorithm by altering its cluster head election and adding fuzzy logic to it. Fuzzy logic is applied where there the clusters formed based on a fixed set of rules are not suitable with their efficiency depending on a number of overlapping metrics. For such applications with uncertainties, this proposed centralized election approach gave outperformed results as compared to the standard probabilistic election approach of LEACH [1]. The fuzzy input variables used by the authors are node energy, concentration and centrality. The drawbacks involve the transmission of more control messages between the nodes and the base station, hence more energy consumption with poor scalability as the second issue.

CHEF [9] consists of another fuzzy cluster head election approach. The proposed distributed clustering algorithm takes residual energy and local distance as the fuzzy input variables. Here, the local distance refers to the sum of distances from all its neighboring nodes within a radius R. The evaluation of the fuzzified values is done using nine fuzzy-if then values. The tentative cluster heads in the defuzzified output have a 'chance' to be elected as the final cluster heads. This chance is computed using the residual energy of the tentative cluster head and its local distance. The energy consumption is affected since the distance between the tentative cluster heads and the base stations is not used for chance computation. The algorithm is observed to have poor scalability results.

A variation of CHEF is proposed by the name EAUCF[10,11] taking residual energy and distance to the base station as the fuzzy input variables. The evaluation of fuzified output is similar to CHEF. The algorithm takes competition radius as the fuzzy output variable and the tentative cluster heads compete to be elected as the final cluster heads based on this fuzzy output variable. The tentative cluster head node having a lower competition radius than the remaining cluster head nodes withdraws from the competition. The output is defuzzified based on the Centre Of Area (COA) approach. Both the CHEF and EAUCF approaches suffer from a low overall performance because of their probabilistic approach of cluster head election and involve a high intra cluster communication thereby affecting the overall energy consumption.

Recently, DUCF algorithm has been proposed by Baranidharan and Santhi [12],. The authors realized the need of a distributed fuzzy unequal clustering algorithm capable enough to eliminate the drawbacks of low overall energy consumption and bring about stability in the network. They focused on First Node Die (FND) and Half Node Die (HND) events for increasing the stability of the network and formation of unequal for balancing the energy consumption. The FND and HND events are considered because due to overloaded work, the events can occur anytime bringing about a huge setback to the performance of the algorithm. The fuzzy input variables used by the authors are the residual energy, node degree and distance to the base station with chance and size as the output fuzzy parameters. The approaches used for the fuzzy inference and defuzzification are the Mamdani and Centroid approaches respectively. The algorithm sets a limit to the number of nodes that can be part of a cluster and also to the size of the cluster nearer to the base station since they act as a router to the other clusters. The performance of the algorithm when compared to the LEACH [1], CHEF [9] and EAUCF [10,11] algorithms gave outperformed results.

V. Conclusion

WSNs being a centre of attention for many ongoing researches have certain constraints to follow. These include effective transmission of information from the source to the sink using the thousands of nodes participating in the network provided that the energy consumed by the nodes is minimized and the network lifetime is prolonged. Also minimum number of nodes can provide a maximum coverage of the region of interest. Two ways of incorporating the above are that the nodes should be limited to doing only the sensing task and sending the information to the cluster heads which then do the remaining work. The process is known as clustering and provides significant advantages over the ordinary functioning of WSNs. Due to the number of advantages observed; many researches are favoring the use of cluster heads in WSNs. This paper gives a brief overview of the developments taken place in this direction with time. Traditional as well as latest contributions discussed will help the reader in determining clustering in WSNs has been improved till date.

References

[1] A.P. Chandrakasan, A.C. Smith, W.B. Heinzelman, "An application specific protocol architecture for wireless micro sensor networks", Transactions on Wireless Communications, Vol. 1, Issue 4, pp. 660–669, 2004.

[2] S. Lindsey, C.S. Raghavendra, "PEGASIS: power-efficient gathering in sensor information systems", Proceedings of the IEEE Aerospace Conference, pp.1125–1130, 2002.

[3] O. Younis, S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks", IEEE Transactions on Mobile Computing, Vol. 3, Issue 4, pp. 660–669, 2004.

[4] M. Ye, C.F. Li, G.H. Chen, J. Wu, "EECS: an energy efficient clustering scheme in wireless sensor networks", Proceedings of the 24th IEEE International Performance, Computing, and Communication Conference IPCCC, pp. 535–540, 2005.

[5] C.F. Li, M. Ye, G.H. Chen, J. Wu, "An energy-efficient unequal clustering mechanism for wireless sensor network", Proceedings of the IEEE International Conference on Mobile Adhoc and Sensor Systems Conference, pp. 596–640. 2005.

[6] H. Li, Y. Liu, W. Chen, W. Jia, B. Li, J. Xiong, "COCA: constructing optimal clustering architecture to maximize sensor network lifetime", Comput. Commun., Vol. 36, pp. 256–268, 2013.

[7] Y. Liao, H. Qi, W. Li, "Load-balanced clustering algorithm with distributed self organization for wireless sensor networks", IEEE Sens. Journal, Vol. 13, Issue 5, pp. 1498–1506, 2013.

[8] I. Gupta, D. Riordan, S. Sampalli, " Cluster-head election using fuzzy logic for wireless sensor networks", Communication Networks and Services Research Conference, pp. 255–260, 2005.

[9] J. Kim, S. Park, Y. Han, T. Chung, "CHEF: cluster head election mechanism using fuzzy logic in wireless sensor networks", Proceedings of the 10th International Conference on Advanced Communication Technology, pp. 654–659. 2008.

[10] H. Bagci and A. Yazici, "An energy aware fuzzy unequal clustering algorithm for wireless sensor networks", Proceedings of the IEEE International Conference on Fuzzy Systems, 2010.

[11] H. Bagci, A. Yazici, "An energy aware fuzzy approach to unequal clustering in wireless sensor networks", Applied Soft Computing, Vol. 13, Issue 4, pp. 1741-1749, 2013.

[12] B. Baranidharan and B. Santhi, "DUCF: Distributed load balancing Unequal Clustering in wireless sensor networks using Fuzzy approach", Applied Soft Computing, Vol. 40, pp. 495–506, 2016.