Open access Journal International Journal of Emerging Trends in Science and Technology

Clustering Techniques for Wireless Sensor Networks

Authors
Sudhir Kumar Gupta[1], Ashish Ranjan[2], Gagandeep Sharma[3]
[1]Keshav Mahavidyalaya, University of Delhi
Email: *Cs.sudhirg@gmail.com*
[2]M.Tech, NIT Patna
Email: *Ashishranjan28@gmail.com*
[3]Aryabhatta College, University of Delhi
Email: *gagandeepsharma@live.com*

Abstract
Wireless sensor network has been emerged as the one of the most promising area of research. The use is not limited to space research, transport, battle fields, surveillance, military, health but many more. The large number of deployed nodes need to be efficiently organized to fulfil the need of data aggregation , load balancing and prolonging network life time Sensor nodes are small power devices so to maintain its long life proper measure is required. Clustering serves this requirement by reducing energy consumption and increasing network scalability. In this article, we highlight the objective of clustering, discuss the different clustering approaches and its classification and finally provide a comparative study of different proposals.
Keywords: *Base Station (BS), Cell Header (CH), inter-cluster communication, intra- cluster communication.*

1. Introduction

Wireless sensor network has been emerged as the most fascinating technology due to its advanced use in different phenomenon of life like smart homes, research, military, space etc. A collection of a large number of wireless sensor nodes is called as WSN in which there is one sink and one sensor field. The sensor nodes are low power, low cost devices which can communicate with other nodes in their radio range. The nodes can perform specific tasks like sensing, tracking and alerting, with which WSNs becomes able to monitor environmental changes, natural phenomena [2], estimating traffic flows, controlling security, monitoring military application [3], and tracking friendly and enemy forces in the battlefields. Network is portioned into a number of small groups, called as cluster, to support data aggregation through efficient network organization. Cluster is a group of sensor nodes in which one node will act as a cluster head, and remaining nodes will act as member nodes. The groups can be overlapped or non-overlapped. The member nodes send their data to the respective CHs where data is aggregated by CHs and sent to the central base station directly or via other CHs in the system.

Component of wireless sensor networks-

The important components of wireless sensor network are discussed below:

Sensor Node: Sensor nodes are electronic devices composed of battery, DAC, actuators processing and communication unit. This is the key element of sensor network which perform multiple tasks like-sensing, data storage, communication and data processing.

Clusters: Large sensor network is partitioned into smaller groups of nodes to make energy efficient and better communication. These small organizational units are called cluster.

Cluster Heads*:* In each cluster, there is one cluster head which perform task of data aggregation and decide schedule of communication.

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.

End User: The data in a sensor network can be used for a wide-range of applications. [1] Therefore, a particular application may make use of the network data over the internet, using a PDA, or even a desktop computer. In a queried sensor network (where the required data is gathered from a query sent through the network). This query is generated by the end user.

Clustering Objectives

Clustering algorithms in the literature varies in their objectives. Different clustering objective are set to fulfil the applications requirements. The most common objective of network clustering is summarized as following [5]:

Load Balancing: The main goal of clustering is to balance the load among them to meet the expected performance[9]. Load balancing becomes more crucial issue in WSNs where CHs are selected from the available sensors [10]. In this scenario equal size of cluster are set to extend network lifetime as it prevents the exhaustion of energy. However distribution of sensors can also influence data delay.

Fault-tolerance: In many applications nodes are susceptible of malfunctioning and physical damage. To avoid loss of important data it is required to have some measure to tolerate the CHs failure. Most important technique is re-clustering of the network.

Increased Connectivity and Reduced Delay: Cluster heads are required to have connectivity to base station. With CHs having very long communication capability, there is no problem but when selected from deployed nodes it require availability of communication path from each cluster head to base station [11] or length of maximum length path should be bounded [12]. Delay is usually factored in by setting a maximum number of hops ‘‘K’’ allowed on a data path. K-hop clustering is K-dominating set problem [13].

Minimal Cluster Count: For the networks where CHs are specialized resource-rich nodes, this scheme is common [14]. Network designer often likes to employ the least number of these nodes since they tend to be more expensive and vulnerable than sensors. The limitation can be due to the complexity of deploying these types of nodes, e.g. when the WSN is to operate in a combat zone or a forest. In addition, the size of these nodes tends to be significantly larger than sensors, which makes them easily detectable which is not preferred in many application including military and border protection.

Maximal Network Longevity: Sensor nodes are energy-constrained. Hence in the harsh environment lifetime of networks becomes a major issue. For the situation where the CHs are richer in resources than sensors it is good to minimize the energy for intra-cluster communication. If possible, CHs should be placed close to most of the sensors in its clusters [13]. In the case of regular sensors, lifetime can be extended by limiting their load. Adaptive clustering, combined clustering and route setup [15] can also be a viable choice to maximize network lifetime.

2. Classification of Clustering Algorithms

3.1 Heuristic Algorithms

A heuristic algorithm has mainly one or both below goals-

To find an algorithm whose rum time is reasonable, however time consumed in cluster set-up is affordable. and/or

To find an optimal solution. These types of algorithms do not depend on any particular metric and hence results provide very good performance.

The main heuristic algorithms for choosing cluster-heads are following-

Linked Cluster Algorithm *(LCA)* [2]: LCA was initially developed for wired sensors but later implemented for wireless sensor networks. In this approach a unique ID number is assigned to each node and nodes can become a cluster head in two way. Node with highest ID among its neighbours becomes cluster head and second, it can become cluster head if there is no cluster head in its neighbours.

Linked Cluster Algorithm 2 *(LCA2)* [3]: This algorithm was proposed to overcome the problem of choosing more number of cluster heads as in LCA. In this approach two types of nodes were categorized- covered and non-covered. All the neighbours of cluster heads are called as covered. A node with lowest ID from non-covered nodes is elected as cluster-head.

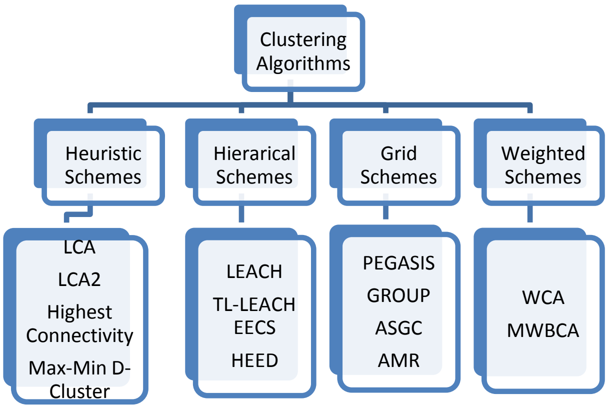

Fig 1: Classification of proposed clustering schemes

Highest-Connectivity Cluster Algorithm [4]: Each node broadcasts total number of its neighbouring nodes. The node with highest connectivity is considered to be cluster head. However in case of tie, node with smaller ID is preferred.

Max-Min D-Cluster Algorithm: In this scheme the authors[11] proposed cluster head selection procedure for distributed scenario in which no node is far from cluster head greater than d hopes. Each node initiates 2d rounds of flooding where first d round, called flood-max, is used for propagation of largest node IDs. In the second d round of flooding, called flood-min is used for smaller nodes to reclaim their territory. Now, each node evaluate the logged in entries according to the following rules

Rule 1: Each node search for its node id in the second d round of flooding. If it is found, it declares itself as cluster head. Otherwise it proceeds to Rule 2.

Rule 2: Node searches for node pairs. Minimum node pair is selected as cluster head. In the absence of node pair, proceed to rule 3.

Rule 3: Nodes selects maximum node id in first d round as the cluster head.

This algorithm requires the following assumptions to be made-:

- Assumption 1: No node id should be propagated farther than d-hops during flooding.
- Assumption 2: All the survived nodes are elected by this method

3.2 Weighted Schemes

Weighted Clustering Algorithm (WCA) *[4]*:

This is demand based, non-periodic, cluster head election procedure[4]. It tries to find most stable configuration in its first set-up. Whenever a sensor node loses its connectivity, new configuration in the network started leading to formation of new topology. This algorithm uses several system parameters such as: the ideal node degree; transmission power; mobility; and the remaining energy of the nodes as input is selection of cluster head and topology. Any or all of the parameters can be used as metric as per application requirement. It is fully distributed in the sense that the entire node perform same responsibility when act as cluster head.

Cluster-head election procedure: The election procedure uses a global parameter called -combined weight, as described in [4]:

$Wv = w1\Delta v + w2Dv + w3Mv + w4Pv$

Where, w1, w2, w3, w4 are the weighing factors for the corresponding system parameters.

3.3 Hierarchical Schemes

LEACH [3] (Low-Energy Adaptive Clustering Hierarchy): The decision for a node to be a cluster head is determined dynamically after some interval of time. The selection depends on priori determined optimal cluster heads and knowledge of how often and last time any specific node has been selected as cluster head. This threshold function is defined as-

(n) ={P/1−P(r mod 1/P), if n € G Otherwise 0 where n is the given node, P is the a priori probability of a node being elected as a cluster-head, r is the current round number and G is the set of nodes that have not been elected as cluster-heads in the last 1/P rounds. Each node during cluster-head selection will produce a random number between 0 and 1. If the number is less than the threshold (T(n)), the node will become a cluster-head.

TL-LEACH [3]: It is an extension of LACH and uses two level of hierarchy. There are two levels of cluster heads in this scheme in addition of normal nodes. In this algorithm, the primary cluster head of each cluster communicates with the secondary and the corresponding secondary communicate with the other nodes in their sub-cluster

EECS (Energy Efficient Clustering Scheme): In this approach cluster-head candidates compete to become cluster-head for a given round. Each node broadcast their residual energy to neighbouring candidates. The node which does not find a node with more residual energy becomes a cluster-head.

***HEED* (Hybrid Energy-Efficient Distributed Clustering):** It is multi-hop clustering algorithm focused on efficient clustering by proper selection of cluster-heads considering physical distance between nodes. The main objectives of HEED are distributing energy consumption to prolong network lifetime, minimizing energy during the cluster-head selection phase, minimizing the control overhead of the network. Cluster heads are determined by two factors- residual energy of each node on probabilistic approach and inter-cluster communication cost.

3.4 Grid Schemes

PEGASIS (Power-Efficient Gathering in Sensor Information Systems): It is a data-gathering algorithm that directly considers energy saving form node instead of cluster formation. The algorithm presents the notion that if nodes form a chain from source to sink, only one node in any given time-frame will transmit to the base station. Data-fusion occurs at every node in the sensor network allowing for all relevant information to permeate across the network [3]. The average transmission range required are much less than LEACH, resulting improved energy consumption than hierarchical clustering approach.

GROUP*:* In this algorithm primary sink dynamically and randomly forms the cluster grid. Queries are forwarded from source to sink are forwarded heads are arranged in a grid like structure. Data queries are forwarded from sink to source is forwarded via Grid Seed (GS) to its cluster-heads, and so on. GS is a node within a given radius from primary sink. Primary sink selects a GS on the basis of residual energy. GS selects cluster head at a distance R along corner of the grid. Each newly selected cluster head then selects more cluster heads based on residual energy along the grid corner.

Conclusions

In this paper we have surveyed the current state of some of most prominent proposed clustering protocols, especially considering power and reliability. In wireless sensor networks, the energy limitations of nodes play a crucial role in designing any protocol for implementation [1]. In addition to this Quality of Service metrics such as data loss tolerance, delay, and network lifetime expose reliability issues when designing recovery mechanisms for clustering schemes [11].We have classified all protocols on the basis of stability and energy efficiency of network. We summarized schemes stating their strengths and limitations.

References

1. I.F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, .Wireless sensor networks: a survey.,Computer Networks 38 (4), 2002, 393.422. doi:10.1016/S1389- 1286(01)00302-4
2. I. F. Akyildiz,W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A Survey on Sensor Netowrks,” IEEE Communications Magazine, vol. 40, no. 8,pp. 102–114, Aug 2002
3. S. Meyer and A. Rakotonirainy, “A Survey of Research on Context-Aware Homes,” Workshop on Wearable, Invisible, Context-Aware, Ambient, Pervasive and Ubiquitous Computing, Adelaide Australia, 2003.
4. B. Warneke, M. Last, B. Liebowitz, Kristofer, and S. Pister, “Smart Dust: Communicating with a Cubic- Millimeter Computer,” Computer Magazine, vol. 34, no. 1, pp. 44–51, Jan 2001. doi:10.1109/2.895117
5. A. A. Abbasi and M. Younis. A survey on clustering algorithms for wireless sensor networks. Computer Communication, 30(14-15):2826–2841, 2007. doi:10.1016/j.comcom.2007.05.024
6. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A survey on sensor networks. IEEE Communications Magazine, 2002. doi:10.1109/MCOM.2002.1024422.
7. Schiller, J. Mobile Communications, second edition ed. Addison Wesley, 2003

8. 16 G. Gupta, M. Younis, Load-balanced clustering in wireless sensor networks, in: Proceedings of the International Conference on Communication (ICC 2003), Anchorage, Alaska, May 2003
9. M. Younis, K. Akkaya, A. Kunjithapatham, Optimization of task allocation in a cluster–based sensor network, in: Proceedings of the 8th IEEE Symposium on Computers and Communications (ISCC'2003), Antalya, Turkey, June 2003.
10. O. Younis, S. Fahmy, HEED: A Hybrid, Energy-Efficient, Distributed clustering approach for Ad Hoc sensor networks, IEEE Transactions on Mobile Computing 3 (4) (2004) 366–379
11. G. Gupta, M. Younis, Fault-tolerant clustering of wireless sensor networks, in: Proceedings of the IEEE Wireless Communication and Networks Conference (WCNC 2003), New Orleans, Louisiana, March 2003.
12. 40 F. Dai, J. Wu, On constructing k-connected k-dominating set in wireless networks, in: Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS'05), Denver, Colorado, April 2005.
13. F. Garcia, J. Solano, I. Stojmenovic, Connectivity based k-hop clustering in wireless networks, Telecommunication Systems 22 (1) (2003) 205–220.
14. E. IlkerOyman, CemErsoy, Multiple sink network design problem in large scale wireless sensor networks, in: Proceedings of the IEEE International Conference on Communications (ICC 2004), Paris, June 2004.
15. K. Dasgupta, M. Kukreja, K. Kalpakis, Topology-aware placement and role assignment for energy-efficient information gathering in sensor networks, in: Proceedings of 8th IEEE Symposium on Computers and Communication (ISCC'03), Kemer-Antalya, Turkey, July 2003.