

A Survey on Clustering Techniques and Load Balancing Algorithms

K.Sundaramoorthy
*Research Scholar,
St.Peters University*

G.Kaviyarasi
Agni College of Technology

Dr.S.Srinivasa Rao Madhane
Adhiparasakthi College of Engineering, Kalavai

Abstract- In this paper, we present a detailed survey on various techniques used for Cluster formation and Load Balancing algorithms to improve the lifetime of the WSN so that it can be extended to reasonable times. There are many reasons where energy will be drained rapidly. Some situations like improper formation of a cluster and also due to traffic, the energy will be consumed more and lifetime of the sensor node will be limited. So, we present a systematic and comprehensive taxonomy for the selection of cluster head, which are discussed in subsequent sections. Special attention was also given to load balancing algorithms which have not yet obtained a wide attention in the literature. Finally, we conclude the paper with insights for research directions about energy conservation in WSNs.

Keywords: Wireless Sensor Network (WSN), Clustering, Load Balancing, Traffic Distribution.

I. INTRODUCTION

Wireless sensor network is a wireless network, which consists of thousands and thousands of sensor nodes that are densely deployed in the required area to monitor the status of the environmental conditions such as weather, earthquake, forest-fire, etc. In real world applications, WSN are used in military applications, positioning and tracking, localization of sensor nodes, etc. These sensor nodes used to communicate with their neighbour nodes using radio signals. Now-a-days, WSN plays vital role in the industry applications like industrial process monitoring, control machine and in the civilian applications like health care monitoring, habitat monitoring, home automation and traffic control.

Even though there are huge advantages of using sensor nodes in various applications, we could also find some disadvantages since, the sensor nodes are smaller in size and contains non-rechargeable batteries. This leads to battery constraint and reduces the lifetime of the network. They also have limited processing speed, the capacity of sensor node is less and also communication bandwidth is a problem.

In this paper, a survey is made on various clustering algorithm and load balancing algorithm to prolong the network lifetime of the sensor nodes. In general the sensor nodes are grouped together and formed as clusters. Each cluster will have a Cluster head (CH) which collects all the scanned information from the nodes attached to it and then the collected information are fused together and the required information is send to sink node i.e. Base station.

Here we classify our paper into five sections. In Section 1, a brief introduction about the node's lifetime. Section 2 tells about the related works. Section 3 is about the various clustering techniques, their advantages and disadvantages. In Section 4, we discussed about different load balancing algorithms used for WSN. Finally, Section 5 ends with conclusion of the paper and the future works that can be done.

II RELATED WORKS

In sensor networks, to increase the lifetime of the node is a great challenge for the researcher. Lot of papers have been published to minimize the energy utilisation of the node in-order to extend the life of the node. But there are

various reasons for the energy loss in the node. To get better performance, the sensor nodes are formed as small cluster using their physical proximity. In every cluster, a cluster head is elected for the coordination of all nodes in that cluster. Due to low power transmission, energy loss at each nodes, etc. there are certain issues in designing WSN.

The algorithm proposed by M.S.Zaman and G.R.Murthy is about the cluster formation and multi path routing method that is used in WSN. The methods like flooding, clustering, leveling, disjoint multi path routing are supported by this algorithm. These method support load balancing and distributes the traffic to every node without any loss of data. This algorithm results in finding the total number of packets that are transmitted from source node to sink node during the process of flooding and the average amount of energy consumed by each node in the every cluster. Hsu-yang, Yunxu proposed cluster based protocol for multipath routing in WSN. Ying Liao, Huan Qi, Weiqun Li [7] proposed more stable and reasonable cluster structure for load balancing clustering algorithm with distributed self organization for WSN, and also improves the network lifetime.

III CLUSTERING TECHNIQUES

In the recent years, Researchers have found clustering of nodes will reduce the power consumption. Therefore to form a cluster between different nodes and also to choose cluster head, there are different clustering schemes available which we have a detailed look. The different schemes of clustering schemes are at a glance, they are Heuristic Schemes, Weighted Schemes, Hierarchical Schemes and Grid Schemes.

a) Heuristic Schemes

These Algorithms will provide nearly the right answer or provide a solution for some instance and not for all instances of the problem. The different algorithm which works in the concept of heuristic schemes are Hill Climbing Techniques, Divide and Conquer algorithms, Branch and Bound Techniques, Dynamic programming, Greedy Techniques[4], Genetic algorithms and k-means clustering algorithm[5]. These heuristic approach aims at balancing power consumption among sensor nodes in order to prolong the network lifetime of the nodes. The heuristic algorithm is based on following goals for formation of cluster and also to obtain optimal solution -

- It selects an algorithm based on the time that is needed to set up clusters and it also finds an optimal solution.
- To choose a cluster head, there are different algorithms available. They are Linked Cluster Algorithm (LCA) [1], Linked Cluster Algorithm 2 (LCA2) [2], Highest-Connectivity Cluster Algorithm [3], Max-Min D-Cluster Algorithm etc.
- b) Weighted Schemes

In this schemes, Weighted Clustering Algorithm (WCA) is a combination of weight metric which includes several system parameters such as node speed, node degree, existence with respect to a node's neighbours, transmission power; mobility, battery life, the time spent as a cluster head, and the remaining energy of the nodes[7]. Depending on the application, one or more metrics will be considered to select the cluster heads. In weighted schemes nodes calculates its weight value and broadcast to all its neighbouring nodes. A node with highest weight among its neighbours will be chosen as Cluster Head; otherwise, it will join to the neighbouring cluster. The main advantage of this scheme is when there is connection loss between the sensor node and cluster head, a new clustering topology is invoked. This will help in power saving i.e. re-election of cluster head occurs less frequently which results in less energy consuming. One more advantage of this algorithm is that it restricts the number of nodes in the cluster. Therefore it is called as good load balancing and also fully distributed algorithm. Though the algorithm provides advantages there are more disadvantages. During the first cluster head selection, the node has to wait for a long time to receive the responses from all its neighbours to make its own decision which increases the latency and overhead[6].

c) Hierarchical Schemes

Hierarchical schemes can be used for homogeneous network or for heterogeneous network. Normally Heterogeneous networks are two level networks, where Cluster head at the first level and Sensor nodes are in the next level. Some algorithm which works in the principle of hierarchical schemes are LEACH, HEED and EECS. The selection of clusters and assigning some tasks to cluster heads can affect the lifetime, scalability, and energy efficiency.

LEACH(Low Energy Adaptive Clustering Hierarchy) is a cluster based protocol, which is used in WSN. It uses the clustering concept to distribute the energy consuming of the node by random rotation of cluster heads. This

protocol uses the randomization cluster-head selection technique depending on the energy available by the sensor nodes. The cluster head selection will be depending upon the threshold function. The threshold function will be calculated using the formula

$$T(n) = \begin{cases} P/1 - P(r \bmod 1/P), & \text{if } n \in G \\ \text{Otherwise } 0. \end{cases}$$

During the cluster head selection each node will generate a random number between 0 and 1. The random number generated by the sensor node will be compared with the threshold function. If the value is less than the threshold value then the particular node will be elected as the Cluster Head. In this way, the energy spent by cluster heads could be reduced to prolong network lifetime. Therefore the LEACH protocol improves the energy efficiency and extend the lifetime of the wireless sensor network.

EECS (Energy Efficient Clustering Scheme) Algorithm is a clustering algorithm which elect the cluster head in a different way when compared with the LEACH protocol. Initially each node broadcast its residual energy to its neighbouring node. When the sensor node receive this message it compare with its residual energy and if it is more than the other sensor node then the particular sensor node will be selected as the cluster head. In the same way formation of cluster vary with other protocols. This protocol support dynamic sizing of cluster and not on minimum distance of nodes. In general the cluster that are in long range will require more energy when compared to the cluster which is near to the base station. This protocol overcome the above problem by distributing the energy over the network. This result in better resource usage and the network lifetime will be extended.

HEED (Hybrid Energy Efficient distributed) is a Clustering algorithm which chooses the cluster head with the following parameters like **residual energy and intra cluster communication** cost. It allows single hop and multihop communication. Each sensor node will communicate with single hop and multihop communication takes place between cluster head and base station. To select a cluster head, each node will calculate the residual energy. To form a cluster, intra cluster communication cost is calculated. The cost value is calculated with the given formula

$$CHprob = Cprob * Eresidual / Emax.$$

This probability value should be below the threshold value. This algorithm have lot of advantages like low message overhead, uniform and non-uniform node distribution and balanced cluster. he disadvantages of this protocol are repeated iteration and the algorithm is very complex, nodes have high residual energy in a particular region of the network.

d) Grid Schemes

A data gathering algorithm in Grid scheme is PEGASIS (Power-Efficient GATHERing in Sensor Information Systems) which reveals the concept of energy savings resulting from nodes and by not forming the clusters directly. The idea of this algorithm is that each node will transmit and receive from their close neighbours and they will also make them as the leader to transmit the data to the base station. In PEGASIS nodes will be forming a chain, this chain can be formed by the sensor node themselves starting from some other node. The base station can also make use of this chain to broadcast messages to all the sensor nodes.

IV LOAD BALANCING ALGORITHMS

When a sensor node collects the data it will send the scanned data to the cluster head, Cluster head will use the concept of data fusion and combine the required data and send to the base station. Due to traffic between the cluster head and the base station, power consuming by cluster head will become more. To avoid this we introduce the concept of Load balancing. To balance the load in the network there are different strategies like centralized load balancing, local or global load balancing, static or dynamic load balancing, sender initiated and receiver initiated load balancing. In Centralized load balancing, as the name implies the amount of load that has to be transferred to other sensor node will be decided by the central node. In the Load balancing, by using the local information of the neighboured processor, it decides upon a load transfer, and thereby minimizing remote communications whereas the Global balancing uses some global information is needed to initiate the load balancing. In static Load Balancing, the load will be assigned initially to all the nodes in the network. In dynamic Load balancing, the load can be changed at the run time depending on the number of sensor nodes participated [8]. The last balancing strategy is sender-initiated and receiver-initiated, in this strategies when data are transferred in the network, due to heavy traffic the

data will be stagnated in the intermediate nodes and that particular nodes will be congested. These congested nodes will try to distribute the load to the lightly-loaded nodes, in the same way the receiver try to share the load from the heavily loaded nodes [9].

V. CONCLUSION AND FUTURE WORK

In this paper we presented a detailed survey on different clustering protocols, for wireless sensor network that saves power and prolongs network life. The life span of the sensor node will also depend on the clustering protocol, we have also discussed different load balancing strategies which can be applied for WSN and also load balancing strategies to reduce the load available in the network, which can be distributed to the lightly loaded nodes. Finally we conclude the paper, an efficient clustering algorithm should be chosen for the cluster formation and selection of cluster head. The load balance algorithm should be used in the network to distribute the load among the neighbored nodes. By using the efficient clustering technique and load balancing algorithm, power consuming in the network can be minimised and lifetime of the network can be increased.

REFERENCES

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, Aug 2002.
- [2] 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.
- [3] 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.
- [4] Nataallia Kokash "An introduction to heuristic algorithms" University of Trento, Italy.
- [5] Jamal N. Al-Karaki Raza Ul-Mustafa Ahmed E. Kamal "Data Aggregation and Routing In Wireless Sensor Networks: Optimal and Heuristic Algorithms"
- [6] M. Chatterjee, S. K. Das, and D. Turgut, "WCA: A Weighted Clustering Algorithm for Mobile Ad hoc Networks", 2002, (5):193–204.
- [7] 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.
- [8] Lee, B. "Dynamic Load Balancing in a Message Passing Virtual Parallel Machine." Technical Report, Division of Computer Engineering, School of Applied Science, Nanyang Technological University, Singapore, 1995.
- [9] Dandamudi, S, "Sensitivity Evaluation of Dynamic Load Sharing in Distributed Systems", Technical Report TR 97 -12, Carleton University, Ottawa, Canada.