A Brief Survey on Clustering and Data Aggregation Routing in WSN

Utkarsh Gupta[1], Prof. Saurabh Kapoor[2]

[1]M.Tech. Scholar, [2]Research Guide

Department of Computer Science and Engg., Gyan Ganga Institute of Technology and Sciences, Jabalpur

Abstract-Nowadays, Wireless Sensor Networks (WSNs) are all over the place and they have turned out to be one of the imaginative innovations that are broadly utilized. They are misused for a large number of utilizations, for example, condition, modern, horticulture, water and sea observing, human services, and so on. Always, WSN is worked of "sensor hubs" from a couple to a few hundreds or even thousands which are in charge of observing a sensor territory and transmit information back to an accumulation point called 'sink'. In this network, every sensor hub is equipped for performing sensory data, preparing and communication with every others in the network without wires.

Keywords-Data aggregation, Wireless Sensor Network, clustering Algorithm, Energy Efficient Aggregation.

I. INTRODUCTION

Wireless sensor networks (WSNs) have been considered as one of the most important technologies used in 21st Century. Basically, a WSN consists of a large number of sensor nodes which are densely deployed over the monitored area in order to collect data about such area. In this Survey, we focus on a specific type of WSNs which called periodic sensor networks (PSNs). In PSN, sensor nodes collect, and then send, data periodically to the end user in order to perform real time data collection for the monitored area. We review a number of PSN applications via some existing examples. Then, we describe challenges faced to PSNs while highlighting the data management challenge as a real problem for such networks.

Wireless sensor network (WSN) research concentrates on working with small, modest, multi-functional sensor nodes that can sense, process, and communicate. WSNs have numerous confinements contrasted with Ad-Hoc networks regarding its sensor nodes' capability of memory storage, processing and the available energy source. These are light weight energy constrained devices that work with little limit DC source. The recharging or replacement of energy sources of the sensor nodes is sometimes difficult or even impractical.

WSNs can be applied to measure humidity, temperature, pollution levels, wind speed and direction, pressure, sound, vibration, and power. With the development of robotized devices and the advancement in wireless communications, it becomes easier to acquire information about the physical environment. Thus, the use of WSN has reduced the challenges met by the conventional method of measuring, processing, and communicating the data to a remote location. In any kind of WSN, these sensor nodes gather and agreeably send this gathered data to a remote base station. The major challenges of the sensor nodes are processing power constraints, battery power limitations, duplicate data gathering, and limited memory power of the network.

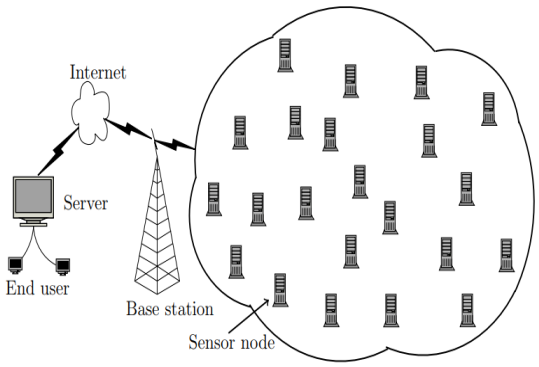

Figure 1: A Wireless Sensor Network

Wireless Sensor Networks (WSNs) have become a highly active research area due to their increasing potential impact on the quality of people's lives. A main application domain where wireless sensor networks are broadly used is environmental data collection and monitoring, where certain conditions or processes need to be monitored periodically, such as the temperature in a conditioned space or pressure in a process pipeline. In such applications, data generated across numerous sensors can produce a significant portion of the big data. Hence, periodic data collection provides two major challenges. First, the network should have a lifetime long enough to fulfill the application requirements. Second, massive and heterogeneous data collected from networks make data management more complex. Researchers' strategies are often targeted to minimize the amount of data collected/communicated by the network without considerable loss in fidelity/accuracy. The goal of this reduction is first to increase the network lifetime, by

optimizing energy consumption of the limited battery for each sensor node, and then to help in analyzing data and making decision. Subsequently in periodic monitoring, the dynamics of the monitored condition or process can slow down or speed up; if the sensor node can adapt its sampling rates to the changing dynamics of the condition or process, over-sampling can be minimized and power efficiency of the overall network system can be further improved. Therefore, in order to keeps the network operating for long time, adaptive sampling approach to periodic data collection constitutes a fundamental mechanism for energy optimization and data reduction.

II. Data Aggregation in Wireless Sensor Network

Data aggregation is defined as the process of summarizing and combining sensor data in order to reduce the amount of data transmission in the network. With the aim of reducing power consumption, data aggregation is the global process of gathering and routing information through a multi-hop network and processing data at intermediate nodes. It attempts to collect the most critical and important data from the sensors nodes and make it available to the Base Station in an energy efficient manner with minimum data latency and minimum possible bandwidth.

Without the use of data aggregation in a WSN, sensor nodes will report all the raw data 2 to the sink. While these data tend to be redundant or correlated, leading to several drawbacks: 1) the redundant data is no sense for the application, 2) the chances of network congestion increase dramatically, 3) the network capacity is wasted, 4) energy consumption increases correspondingly. By previous studies, temporal and spatial correlations are often based on the raw data. For a given sensor node, temporal correlation exists in the data collected at different time instants, while spatial correlation occurs when the data is collected from the neighboring sensor nodes.

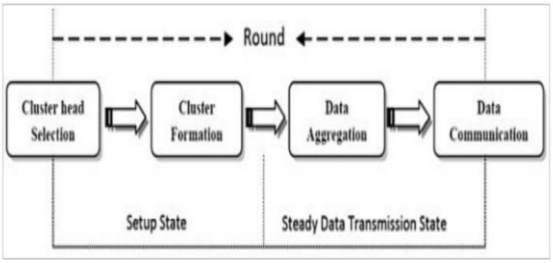

Fig. 2: Block Diagram of Data Aggregation in WSN.

- *Energy Clustering in Wireless Sensor Networks*

In clustering, the sensor nodes are partitioned into different clusters. Each cluster is managed by a node referred as cluster head (CH) and other nodes are referred as cluster nodes. Cluster nodes do not communicate directly with the sink node. They have to pass the collected data to the cluster head. Cluster head will aggregate the data, received from cluster nodes and transmits it to the base station. Thus minimizes the energy consumption and number of messages communicated to base station. Also number of active nodes in communication is reduced. Ultimate of clustering the sensor nodes is prolonged network lifetime.

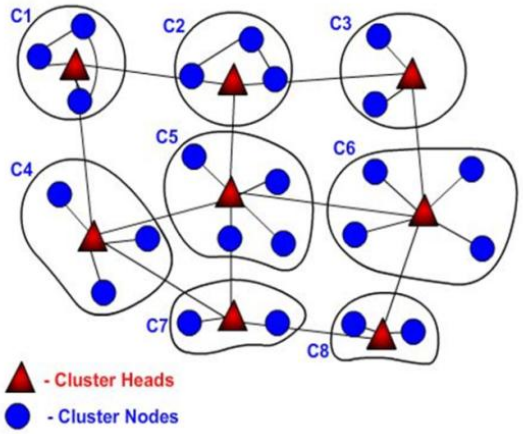

Fig. 3 Clustered Sensor Network.

Sensor Node: It is the core component of wireless sensor network. It has the capability of sensing, processing, routing, etc.

Cluster Head: The Cluster head (CH) is considered as a leader for that specific cluster. And it is responsible for different activities carried out in the cluster, such as data aggregation, data transmission to base station, scheduling in the cluster, etc.

Base Station: Base station is considered as a main data collection node for the entire sensor network. It is the bridge (via communication link) between the sensor network and the end user. Normally this node is considered as a node with no power constraints.

Cluster: It is the organizational unit of the network, created to simplify the communication in the sensor network.

- *Energy Efficient Data Aggregation*

The most energy consumable operation is transmitting data by a sensor node, comparing it with the energy consumption of in-network computation which is negligible. The energy trade-off between communication and computation provides applications benefit when processing the data at the network side rather than simply transmitting sensor data. In this study, energy efficient two phase data aggregation technique for clustering based PSN.

In WSN, sensor node lifetime is highly related to the power consumption of its battery as it is the low source of

energy, and it is difficult and cost ineffective to recharge it in most cases. Sensor nodes use their limited energy in computation and transmission processes in a wireless environment, but the power consumption is at the highest level when sending and receiving messages.

Data aggregation and data reduction approaches are proposed to conserve energy in WSN by reducing the amount of data sent from sensor nodes to their appropriate sink. To save overall energy resources, sensing data are aggregated along the route from sensors to sink. In addition, the amount of data generated in large sensor networks is usually redundant which makes the data aggregation methods essential to eliminate redundant transmissions. It is important to highlight that data aggregation, by eliminating redundant data, should not affect the quality of data.

III. LITERATURE SURVEY

Sr. no.	Title	Author	Year	Approach
1	Clustering and Data Aggregation in Wireless Sensor Networks Using Machine Learning Algorithms	S. K and V. Vaidehi	2018	This paper presents a literature review of different machine learning based methods which are used for clustering and data aggregation in WSN and proposes an improved similarity based clustering and data aggregation.
2	Density, distance and energy based clustering algorithm for data aggregation in wireless sensor networks	H. Lin, R. Xie and L. Wei,	2017	In this paper, we propose a new clustering method called Density, Distance and Energy based Clustering (DDEC) to improve network performance. DDEC partitions the network into clusters with similar member number, so as to achieve load balancing.
3	Performance analysis of the energy efficient clustering models in wireless sensor networks	S. Vançin and E. Erdem	2017	In this paper, SEED (sleep-wake energy balanced distributed) algorithm was compared with both clustering methods such as, LEACH, mod-LEACH and PEGASIS for homogeneous sensor networks and SEP, DEEC and CEEC for heterogeneous sensor network.
4	Data clustering in wireless sensor network implemented on self organization feature map (SOFM) neural network.	M. Mittal and K. Kumar	2016	Wireless sensor network is one of the most promising communication networks for monitoring remote environmental areas. In this network, all the sensor nodes are communicated with each other via radio signals. The sensor nodes have capability of sensing, data storage and processing.
5	Clustering technique for Wireless Sensor Network	K. Desai and K. Rana	2015	In our proposed algorithm cluster head (CH) selection is carried out using distance between nodes and energy of the nodes. They are selected in such a way that, the distance between them is maximum and energy of the selected CH is more.
6	Design and implementation of high event density area centered clustering based routing	B. Jeon, B. Kang and S. Park	2012	In this paper, we propose high event density area centered clustering based routing (HEDACR) for efficient energy consumption in WSN. HEDACR forms a high density tree based cluster centered at the high event density area.

S. K and V. Vaidehi [1] Wireless Sensor Networks (WSN) are resource constrained. Clustering and data aggregations are used to reduce the energy consumption in the network by decreasing the amount of data transmission. Machine Learning algorithms such as swarm intelligence, reinforcement learning, neural networks significantly reduce the amount of data transmission and use the distributive characteristics of the network. It provides a comparative analysis of the performance of different methods to help the designers for designing appropriate machine learning based solutions for clustering and data aggregation applications. This paper presents a literature review of different machine learning based methods which are used for clustering and data aggregation in WSN and proposes an improved similarity based clustering and data aggregation, which uses Independent Component Analysis (ICA).

H. Lin, R. Xie and L. Wei, [2] Wireless sensor networks (WSNs) are wireless networks which consist of distributed sensor nodes monitoring physical and environmental conditions. Due to the energy limit of sensor nodes, prolonging lifetime of wireless sensor networks (WSNs) is a big challenge. In this paper, we propose a new clustering method called Density, Distance and Energy based Clustering (DDEC) to improve network performance. DDEC partitions the network into clusters with similar member number, so as to achieve load balancing. Then a cluster head is selected for each cluster based on three criteria: residual energy, distance and density, which achieves to minimize intra-communication cost and prolong cluster lifetime. In our performance analysis, we compare DDEC with another clustering method called DDCHS. The results show that DDEC outperforms DDCHS in terms of alive node number and energy consumption.

S. Vançin and E. Erdem [3] In wireless sensor networks, energy-oriented routing and data aggregation methods are being developed owing to the fact that sensor nodes have limited energy and memory capacity. Cluster-based heterogeneous routing protocols, a popular part of routing technology, have proven effective in topology management, energy consumption, data collection or fusion, reliability, or stability in a distributed sensor network. In this paper, SEED (sleep-wake energy balanced distributed) algorithm was compared with both clustering methods such as, LEACH, mod-LEACH and PEGASIS for homogeneous sensor networks and SEP, DEEC and CEEC for heterogeneous sensor networks in terms of the number of alive nodes in the network and the number of packets sent to the base station (BS) as two different scenarios. Simulation results clearly demonstrate that the SEED algorithm is more robust than the other three methods and increases the network efficiency.

M. Mittal and K. Kumar [4] Wireless sensor network is one of the most promising communication networks for monitoring remote environmental areas. In this network, all the sensor nodes are communicated with each other via radio signals. The sensor nodes have capability of sensing, data storage and processing. The sensor nodes collect the information through neighboring nodes to particular node. The data collection and processing is done by data aggregation techniques. For the data aggregation in sensor network, clustering technique is implemented in the sensor network by implementing self organizing feature map (SOFM) neural network. Some of the sensor nodes are selected as cluster head nodes. The information aggregated to cluster head nodes from non cluster head nodes and then this information is transferred to base station (or sink nodes). The aim of this paper is to manage the huge amount of data with the help of SOM neural network. Clustered data is selected to transfer to base station instead of whole information aggregated at cluster head nodes. This reduces the battery consumption over the huge data management. The network lifetime is enhanced at a greater extent.

K. Desai and K. Rana [5] Wireless Sensor Network (WSN) consists of small nodes with sensing, computation, and communications capabilities. Sensor node senses the data and sends data to the base station for further processing. These sensor nodes mainly rely upon batteries for energy, which get drained at a quicker rate due to the computation and communication and this decreases the lifetime of the network. In order to solve this problem, various clustering techniques are introduced to improve the lifetime of WSN. Clustering has turned out to be a powerful approach for formulating the network into a linked hierarchy. The main goal of clustering algorithms is to gather and aggregate data in an energy efficient manner so that network lifetime is enhanced. In our proposed algorithm cluster head (CH) selection is carried out using distance between nodes and energy of the nodes. They are selected in such a way that, the distance between them is maximum and energy of the selected CH is more. All the cluster heads send data to the CH located closer to the BS and in turn, it aggregates and transmits data to the Base Station (BS).

B. Jeon, B. Kang and S. Park [6] Energy consumption is the most important issue in a wireless sensor network (WSN). As the WSN scale increases, the required energy consumption and bandwidth are increased. In this paper, we propose high event density area centered clustering based routing (HEDACR) for efficient energy consumption in WSN. HEDACR forms a high density tree based cluster centered at the high event density area. The main idea of HEDACR is to form a multi-hop short distance routing path and for all sensors to participate in the data transmission to diffuse energy consumption for data transmission. We developed the environmental data gathering and aggregation sensor (EGAS) to realize HEDACR. We experimented on HEDACR in a test bed. The result of experiment indicates that HEDACR reduces the energy consumption and extends the network lifetime as compared with LEACH using one-hop.

IV. PROBLEM IDENTIFICATION

Recent improvements in technology provide us cheap and tiny electronic devices with various sensors on it. These tiny devices are called 'sensor nodes' and they have great abilities. The aim of using sensor nodes is to sense the environment and process and/or transfer collected information to an analysis center. Sensor nodes are usually battery powered and their trans- mission range is very low. Therefore, these sensor nodes can establish a network to propagate their data to long distances. Moreover we can

perform our algorithms on more challenging topologies with higher trans- mission ranges. Therefore we can find out effects of obstacles among nodes, such as trees, walls, or elevation differences etc.

V. CONCLUSION

WSNs can be applied to measure humidity, temperature, pollution levels, wind speed and direction, pressure, sound, vibration, and power. With the development of robotized devices and the advancement in wireless communications, it becomes easier to acquire information about the physical environment. Thus, the use of WSN has reduced the challenges met by the conventional method of measuring, processing, and communicating the data to a remote location.

We defined several algorithms with different features, and as a future work, we can combine some of the features of these algorithms to select cluster heads more efficiently. For example, we can select cluster heads from nodes with more neighbors and higher energy levels.

REFERENCES

[1] S. K and V. Vaidehi, "Clustering and Data Aggregation in Wireless Sensor Networks Using Machine Learning Algorithms," 2018 International Conference on Recent Trends in Advance Computing (ICRTAC), Chennai, India, 2018, pp. 109-115.

[2] H. Lin, R. Xie and L. Wei, "Density, distance and energy based clustering algorithm for data aggregation in wireless sensor networks," 2017 IEEE/CIC International Conference on Communications in China (ICCC), Qingdao, 2017, pp. 1-5.

[3] S. Vançin and E. Erdem, "Performance analysis of the energy efficient clustering models in wireless sensor networks," 2017 24th IEEE International Conference on Electronics, Circuits and Systems (ICECS), Batumi, 2017, pp. 247-251.

[4] M. Mittal and K. Kumar, "Data clustering in wireless sensor network implemented on self organization feature map (SOFM) neural network," 2016 International Conference on Computing, Communication and Automation (ICCCA), Noida, 2016, pp. 202-207.

[5] K. Desai and K. Rana, "Clustering technique for Wireless Sensor Network," 2015 1st International Conference on Next Generation Computing Technologies (NGCT), Dehradun, 2015, pp. 223-227.

[6] B. Jeon, B. Kang and S. Park, "Design and implementation of high event density area centered clustering based routing," 2012 International Conference on Computing, Networking and Communications (ICNC), Maui, HI, 2012, pp. 247-251.

[7] H. Lin, L. Wang, and R. Kong, “Energy efficient clustering protocol for large-scale sensor networks,” IEEE Sensors Journal, vol. 15, no. 12, pp. 7150–7160, 2015.

[8] A. Thakkar and K. Kotecha, “Cluster head election for energy and delay constraint applications of wireless sensor network,” IEEE sensors Journal, vol. 14, no. 8, pp. 2658–2664, 2014.