A Survey on: Energy Efficient Big Data in Cluster Based Wireless Sensor Networks.

Mrs. Kalpavi C Y[1] ,Dr.Veena C S[2]

[1]Assistant Professor, Department of ECE AIT, Bangalore.

[2]Professor, Department of ECE SaIT, Bangalore.

Abstract

A wireless sensor network (WSNs) is an important application for real time monitoring and detecting some phenomenon. Clustering is an effective approach to prolong the network lifetime, but many existing clustering algorithms has a static sink node, which will result in a hot spots problem. Big-data is a popular term in the field of information and communication technology. WSN is one of the eminent contributors of big data. WSN contains numerous sensor nodes that cooperatively monitor an environment. Each network consists of sensor node, memory, and communication device. Data generated by single sensor node is small but data generated by distributed sensor network is significantly large and it is termed as big data. The critical issue in WSN is energy consumption and data gathering. WSN is proposed for the risk analysis of industrial operations. Clustering data transmission structure is then established on the basis of the received signal strength indicator (RSSI) and residual energy information.

Keywords: *Big data, Clustering, Distributed Sensors, Residual Energy and Static sink node.*

I. INTRODUCTION

Wireless sensor networks (WSNs) are consisted of many tiny nodes which possess the capacity of communication and storage. These nodes are densely deployed either inside the phenomenon or close to it and can communicate with each other in a multi hop communication mode. Due to self-organizing capabilities, these nodes are usually scattered in the sensing field and do not need artificial maintenance. With the development of the electronics manufacturing technology, the different types of sensors can be applied in military reconnaissance, surveillance, environment monitoring and natural disaster relief, health monitoring, and so on.

Hierarchical protocols include Low-Energy Adaptive Clustering Hierarchy (LEACH), Threshold sensitive Energy Efficient sensor Network protocol (TEEN) and Power-efficient Gathering in Sensor Information Systems (PEGASIS). LEACH is the first implemented cluster-based routing protocol and other cluster-based routing protocol

ISBN : 978-93-87793-86-6

can be viewed as an improvement of LEACH. These protocols aim at selecting a cluster in a small area to transmit data to sink node or base station, which can meet the requirements of the scalability and can control nodes' energy consumption, in this way that the lifetime of the network will be extended. The cluster-based network's topology is shown in Figure 1.

Big data is a term that represents the data having high volume, velocity, variety and veracity. This massive data is stored in cloud environment but it is difficult to capture, store and analyze using the currently available technologies. Big data gathering is illustrated in Figure 2.

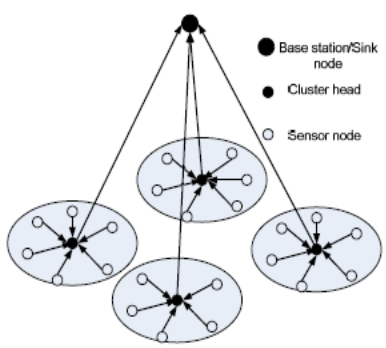

Figure 1. Cluster-based network topology

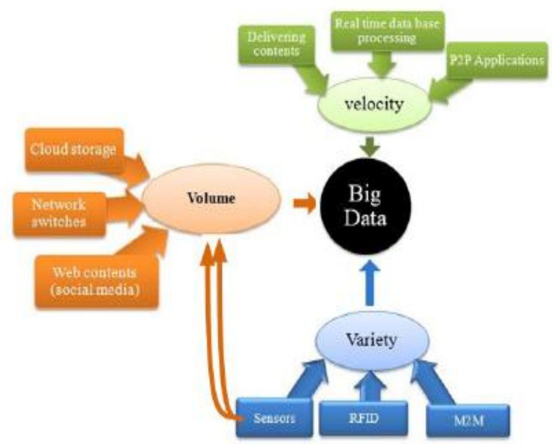

Figure 2: Big Data Gathering.

II. CLUSTER BASED TOPOLOGY (CBT)

Ye et al. and Liu discussed CBT has been widely used in WSNs for data gathering, data dissemination, target tracking, etc. Clustering is a proficient method, widely used in dense WSNs, scalability to hundreds or thousands of nodes. Scalability in this context implies the need for proficient resource management, data fusion/aggregation, load balancing, robustness, etc. It contains set-up phase and steady-state phase. In set-up phase, cluster head election is an important task, which is achieved by a variety of methods like centralized (i.e., assigned by base station) or distributed (i.e., random method, probabilistic, residual energy and election phase) election in intra cluster communication. Thereafter, the entire cluster head connected to the sink with either direct-hop (i.e., cluster head directly connected with sink) or multi-hop (i.e., communication with cluster head to cluster head and then sink) communication techniques in an inter cluster communication. Then, the steady-state phase can be initiated to forward the data packets from the cluster members to sink with inter cluster and intra cluster communication. For mobility-based environments, intra cluster communication with frequent cluster member changes of cluster head and

ISBN : 978-93-87793-86-6

inter cluster communication with multi-hop techniques or direct communications cannot offer a guaranteed data transmission rate. It diminishes the performance of the entire network. E.g., LEACH, HEED, BCDCP, CCS, EAP, HGMR and MBC.

Heinzelman et al. proposed LEACH (Low Energy Adaptive Clustering Hierarchy), one of the most popular clustering algorithms for WSNs, which is absolutely distributed and there is no global knowledge required to the network. The cluster head election and formation is based on the received signal strength and threshold value, and then the data fusion/aggregation operations can be achieved by cluster. Optimal number of cluster heads is equivalent to 5% of the total number of sensor nodes. All the sensor nodes have a chance to act as a cluster head for a particular time, in order to balance the energy dissipation of sensor nodes. It can be achieved by the sensor node selecting a random number between 1 and 0. The sensor nodes die slowly, randomly and dynamic clustering improves the lifetime of the network. On the other hand, LEACH uses single hop transmission (i.e., all the cluster members can directly communicate with the cluster head and forward the data to the sink), it is not applicable to large-scale mobile WSNs. Furthermore, dynamic clustering brings extra overhead (e.g., link failure, frequent cluster head changes and advertisements, etc.), which may increase the energy conservation.

Younis and Fahmy have proposed HEED (Hybrid Energy Efficient Distributed Clustering), an energy efficient multi-hop clustering algorithm, which selects the cluster head based on the high average residual energy when compared to neighbor nodes and intra cluster communication cost. HEED provides uniform CH distribution across the network and load balancing. Multi hop communication between CHs and BS endorse energy conservation and scalability. However, more cluster head generation from tentative cluster head may leave several uncovered senor nodes. These cluster heads may not have cluster members, and that can’t fall on the coverage range of other cluster head. HEED broadcasts lot of control packets, which needs several iterations to form clusters. Cluster head near to the sink, may die earlier since these nodes have more routing packets overhead.

Muruganathan et al. introduced BCDCP (Base-Station Controlled Dynamic Clustering Protocol), a centralized cluster based routing protocol with complex computational capability of the sink. It maintains the uniform distribution of cluster head over the network and nearly identical number of cluster members to balance the cluster head overload. On each round, the sink collects the information of the entire sensor nodes residual energy and computes the average energy level of these nodes in the network to select a set of nodes that are elected as cluster head. It exploits a high-energy sink to form clusters and utilizes the MST (Minimum Spanning Tree) to connect cluster heads. It arbitrarily picks up a cluster head to send the data to the sink. However, there are few disadvantages in BCDCP such as, BCDCP uses centralized algorithm (i.e., one-hop routing scheme) which gives poor scalability and robustness to large-scale networks than distributed algorithms. Each node should update their location and energy level to the sink every round, and it is not an appropriate protocol for periodic data retrieval based reactive networks.

Jung et al.suggested CCS (Concentric Clustering Scheme) which aims to reduce the energy conservation in PEGASIS. The network is partitioned into a mixture of concentric circular tracks which means various cluster node in the network is allocated with its own level. In addition, chains are constructed inside the track as that in PEGASIS. The track adjacent to the BS is level 1 and the level number is incremented with the addition of the distance to the BS. All the sensor nodes in each level collect the data, fuse its own data and then follow the chain to forward the data to the neighbor sensor node on both sides of cluster head. Consequently, all the cluster head in each level forwards the data to the lower cluster head. Finally, level 1 cluster head broadcasts these data to the sink. While compared to PEGASIS, CCS has significant energy savings, the reason being decreasing transmission distance between cluster head and sink. Concentric cluster partition reverses the data flow from the sink reduced, which reduces the energy consumption over the entire network. However, the disadvantages of the protocol are (i) sensor node allocation in each level is unequal, (ii) residual energy is not considered for cluster head selection, (iii) neighbor nodes on long chain may communicate with low radio power may cause large delay, (iv) the cluster head election for next hop is set on the location information rather than the residual energy of sensor nodes which dissipates the energy of cluster head quickly on the path among cluster head and even energy hole will appear in the network.

Liu et al. recommended EAP (Energy Aware Routing Protocol), a variant of LEACH and HEED to enhance the performance of LEACH and HEED. It reduces the energy consumption of prolonged WSNs. EAP is an energy efficient data collection protocol. It groups the sensor nodes for intra cluster communication, and it builds the routing tree between cluster heads to make inter cluster communication. Conversely, EAP is not suitable for mobility-based WSNs, since it cannot provide a stable link whenever the sensor node on mobility. Koutsonikolas et al. have proposed a location based multicast protocol called as HGMR (Hierarchical Geographic Multicast Routing). HGMR offers very simple membership management with free scalability, and energy efficient routing methods on different hierarchies. However, network may be partitioned into set of small cells which may lead to sub-optimal routing paths from the root node to multicast group members, and unstable energy utilization around APs (Access Points).

Deng et al. designed MBC (Mobility Based Clustering protocol), wherein all the sensor nodes have a chance to elect the cluster head based on the threshold value (i.e., residual energy and mobility) of each round. It does intra cluster (i.e., all the cluster members transmit the sensed data to the cluster head in an allocated time slot) and inter cluster communication (i.e., cluster head collects the data from cluster members and forwards it to the sink) in each round.

ISBN : 978-93-87793-86-6

MBC provides better performance than LEACH, HEED and other existing protocols on mobility-based environment, but MBC fails to address the critical node occurrence problem which causes link breakage, packet dropping and reduces the network utilization.

III. AREA BASED APPROACHES

In this category, the MEs visiting different areas for data collection have different effects on the network performance, such as load balance and decreased data gathering latency. Sensor nodes in other areas need to transmit their data to these visited areas via multi-hop communication. The authors of propose an Energy-Balanced Data Collection (EBDC) mechanism. It considers a circular monitoring region, in which sensor nodes in different rings have different data-relay in workloads. To balance the energy consumption, the mobile sink needs to move along different tracks with different number of sweep repetitions to collect data. The number of sweep repetitions in each track is calculated by the energy consumption of each sensor node when the mobile sink move along that track. Similar to, in order to balance the non-uniform energy consumption of sensor nodes in a circular area, the authors of present a clustering-based routing algorithm in wireless sensor networks with mobile sink (CRA). In the CRA, the mobile sink moves around a specified circle and the circular area near this trajectory is defined as buffer area. The buffer area has the largest load in the network, so in order to achieve maximum network lifetime, the main task in the CRA is to determine the radius of the buffer area and the radius of the moving circle of the mobile sink. Moreover, an improved algorithm CRA-1 is also introduced by using density controlling to further balance the network load. A biased sink mobility with adaptive stop times is studied in, as a method for efficient data collection in wireless sensor networks. To achieve a balance of energy consumption and data gathering latency, when selecting the next area to visit, the mobile sink favors less visited areas to cover the network area faster, and adaptively stops more time in some regions that tend to produce more data.

IV. NETWORK AND ENERGY CONSUMPTION MODEL

The WSN for our TCBDGA consists of *N* static and homogeneous sensor nodes deployed randomly over an *L* _*L* square area, as shown in Figure 4. There is a static BS located at the center of the distributed area. A mobile sink moves in the end periodically from the BS to collect data. The mobile sink is assumed to be with unlimited energy, memory and computing resources. All sensor nodes are initially provisioned with the same amount of energy *E*0 and location-aware abilities by using localization technology and the mobile sink knows all nodes' locations. Each node has a limited communication range *Rs*, and communication is always successful if other node is within its transmission range. Each of them generates data with the same rate and sends them only once in each data collection round. We assume that the whole deployment area is full covered and well connected.

ISBN : 978-93-87793-86-6

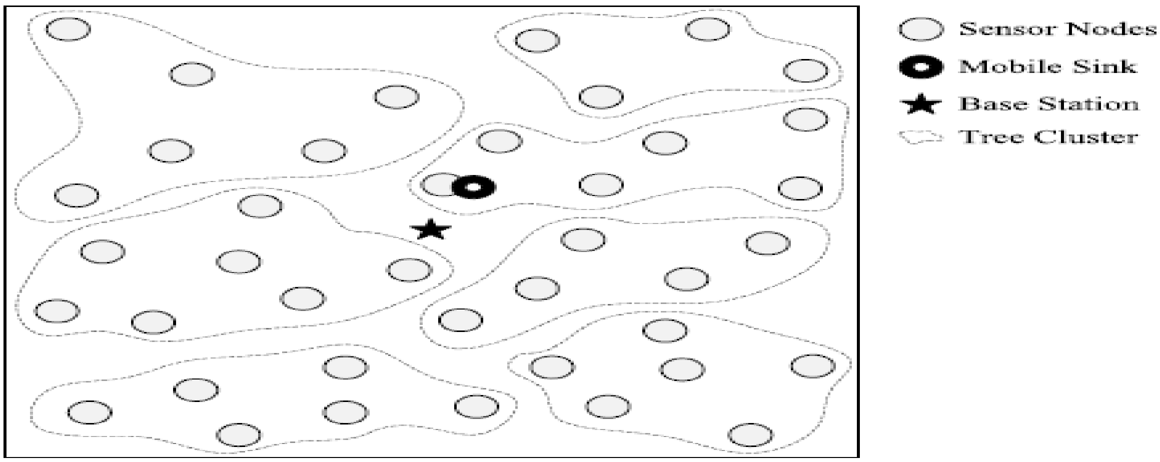

Figure 4: Network Model

TABLE 1. ANALYSIS RESULT

Parameters	Number of Mobile Sinks							
	1	2	3	4	5	6	7	8
Simulation start time (s)	0	0	0	0	0	0	0	0
Data transfer end time (s)	88.18	60.71	51.92	48.19	49.64	46.32	51.70	44.77
Packets sent	2934	1462	1827	1502	1862	1562	1789	1443
Packets received	2933	1461	1826	1502	1862	1556	1788	1442
Number of packets dropped	1	1	1	0	0	6	1	1
PDR Value	0.9996	0.9993	0.9994	1	1	0.9961	0.9994	0.9993
Average remaing energy	16.70	18.30	17.90	18.30	17.90	18.24	17.98	18.37
Total number of nodes	26	27	28	29	30	31	31	33
Number of MS	1	2	3	4	5	6	7	8
Initial total energy	520	540	560	580	600	620	640	660

ISBN : 978-93-87793-86-6

Table 2: Simulation Parameters

Parameters	*Representation*	*Unit*
N	The number of the sensor nodes	100
E_0	Initial energy	0.5J
E_{DA}	Data aggregation energy	5nJ/bit
E_{elec}	Energy dissipation to run the radio device	50nJ/bit
ε_{fs}	Free space model of transmitter amplifier	10 pJ/bit/m ^{2}
ε_{mp}	Multi-path model of transmitter amplifier	0.0013pJ/bit/m^4
l	Packet length	2000bits
d_0	Distance threshold	$\sqrt{\varepsilon_{fs} / \varepsilon_{mp}}\ m$

V.CONCLUSIONS

On considering the raising the impact of WSNs on real-time civil and military applications, more number of sensor nodes are required to monitor the large-scale areas. In this paper, VELCT (Velocity Energy-efficient and Link-Aware Cluster-Tree) is a proficient method is proposed to construct a mobility-based auspicious network management architecture for WSNs, to exploit the network lifetime, connection time, residual energy, RSSI, throughput, PDR and stable link for mobile sensor nodes, whereas each cluster member chooses the cluster head with better connection time and forwards the data packets to the corresponding cluster head in an allocated time slot. Similarly, the sink or DCN elects the one-hop neighbor DCN or cluster head with maximum threshold value, connection time, RSSI and with less network traffic. From the simulation results, it is evident that VELCT provides more stable links, better throughput, energy utilization and PDR with reduced network traffic than EEDCP-TB, CREEC, CTDGA, MBC and CIDT.

VI. REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A survey on sensor networks", *IEEE Communications Magazine, vol. 40, no. 8, 2002, pp. 102-114.*

[2] J. Yick, B. Mukherjee, D. Ghosal, "Wireless sensor network survey", *Journal of Computer Networks, vol. 52, no. 12, 2008, pp. 2292-2330.*

[3] S. Tyagi, N. Kumar, "A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks", *Journal of Network and Computer Applications, vol. 36, no. 2, 2013, pp. 623-645.*

[4] W.R. Heinzelman, A. Chandrakasan, H. Balakrishnan, "Energy-efficient communication protocol for wireless Microsensor Networks", *International Conference on System Sciences, 2000, January.*

[5] A. Manjeshwar, D.P. Agrawal, "TEEN:a protocol for enhanced efficiency in wireless sensor networks", *International Parallel and istributed Processing Symposium, 2001.*

[6] S.Lindsey, C.S. Raghavendra, "PEGASIS:power efficient gathering in sensor information systems", *IEEE Aerospace Conference, 2002.*

[7] R. Jaichandran, A. Irudhayaraj, J. Raja, "Effective strategies and optimal solutions for hot spot problem in wireless sensor networks (WSN)", *International Conference on Information Sciences Signal Processing and their Applications, 2010.*

[8] Y. Shi, Y.T. Hou, "Theoretical Results on Base Station Movement Problem for Sensor Network", *IEEE Conference on Computer communications, 2008.*

[9] J. Luo, J.-P. Hubaux, "Joint Sink Mobility and Routing to Increase the Lifetime of Wireless Sensor Networks: The Case of Constrained Mobility", *ACM/IEEE Transactions on Networking, vol. 18, no. 3, pp. 871 – 884, 2010.*

[10] Y. Sun, Y. He, B. Zhang, X. Liu, "An energy efficiency clustering routing protocol for WSNs in confined area", *Mining Science and Technology (China), vol. 21, no. 6, 2011.*

[11] C.-J. Lin, P.-L. Chou, C.-F. Chou, "HCDD: Hierarchical cluster based data dissemination in wireless sensor networks with mobile sink", *International Conference on Wireless Communications and mobile computing, 2006.*

[12] A. Amis, R. Prakash, T. Vuong, D. Huynh, "Max-min d-cluster formation in wireless ad hoc networks", *IEEE Conference of the Computer and Communications Societies, 2000.*

[13] X. Yuan, R. Zhang, "An energy-efficient mobile sink routing algorithm for wireless sensor networks, in Wireless Communications", *International Conference on Wireless Communications, Networking and Mobile Computing, 2011.*

ISBN : 978-93-87793-86-6

[14] B. Nazir, H. Hasbullah, "Mobile sink based routing protocol (MSRP) for prolonging network Life time in clustered wireless sensor network", *International Conference on Computer Applications and Industrial Electronics, 2010.*

[15] K. Akkaya, M. Younis, "A survey on routing protocols in wireless sensor networks", *Elsevier Ad Hoc Networks, vol. 3, no. 3, 2005, pp. 325-349.*

[16] A.W. Khan, A.H. Abdullah, M.A. Razzaque, J.I. Bangash, "VGDRA: A Virtual Grid-Based Dynamic Routes Adjustment Scheme for Mobile Sink-Based Wireless Sensor Networks", *IEEE Sensors Journal, vol. 15, no. 1, 2015, pp. 526 – 534*