Data Aggregation Techniques Over Wireless Sensor Network- A Review

Shilpa S G
Asst. Professor,
Dept. of Computer Science and Engg,
Don Bosco Institute of Technology, Bangalore.

Dr. S. Meenakshi Sundaram
Professor and Head,
Dept. of Computer Science & Engineering, GSSS Institute of Engineering and Technology for Women, Mysore

***Abstract*—Wireless sensor networks (WSNs) comprise of several sensor nodes which sense and measure the environmental data and report the same to sink or base station. These sensor nodes sense the similar information and sends it to sink node. This leads to redundancy at the sink node. Sink nodes consumes more energy in processing these redundant packets. As the sensor nodes have limited battery power and small data storage most of the battery power is wasted by eliminating redundancy at sink node and shortens the network lifetime. Hence to conserve the energy of a node and to improve the lifetime of WSN, we need a technique to eliminate redundant sensed data and aggregation of similar sensor values to reduce the communication overhead. Data Aggregation is an efficient way to reduce the huge volume of data generated in WSNs by eliminating the redundancy among wireless sensor networks. In this paper, we have focused on importance of data aggregation technique and different data aggregation protocols based on network.**

Keywords: WSN, Data Aggregation, Data Gathering, Cluster Based Protocols.

I. INTRODUCTION

Wireless Sensor Networks consist of densely deployed, low-cost, low-power, multifunctional sensors. In WSNs, the basic function of sensor networks includes collaborative sensing, sampling, computing and broadcasting the sensed information. Sensor nodes have limited battery power which is not possible either to recharge or replenish the node of the battery once it is deployed in the field [1]. In a deployed WSN, the communication of nodes with in their radio range and data gathering consumes more energy. Hence there is a necessary to reduce the energy consumption at each sensor node to increase the network life time of WSN. This can be accomplished by removing redundant information in wireless network. Because most of the energy of a node is wasted during processing the redundant data. Thus eliminating redundancy is one of the solution to improve network lifetime. A sensor node is a tiny device that composed of three portions [2]-

- The data is collected from the physical environment using the sensing subsystem.
- Data manipulation and storage using the processing subsystem.
- Data transmission using wireless communication subsystem.

Information perceived on sensed data can be employed in agriculture and livestock, driving or perhaps in providing security at home or in public areas. A key demand from each the technological and commercial purpose is to supply adequate security capabilities. Fulfilling privacy and security necessities in an appropriate architecture for WSNs giving pervasive services is crucial for user acceptance. 5 key optionshave to be compelled to be thought of once while developing WSN solutions: scalability, security, reliability, self-healing and robustness.

II.DATA AGGREGATION

The nodes which are in same radio range may sense the redundant data and transmits the same to sink node. Then it is a challenging for the sink node to manage such large amount of data. This problem can be solved by a data driven approach called “Data Aggregation”. The approach data aggregation is the power-saving mechanism. It is the process of combining the data coming from various sources and en route them after removing redundancy, such as to improve overall network lifetime. This can significantly help to reduce the consumption by eliminating redundant data [3].

The functionality of data aggregation is performed continuously in order to improve the bandwidth and energy utilization, but it may impact badly on other performance metrics such as delay, accuracy, fault tolerance, etc. However the objective of the data aggregation is to eliminate the redundant data transmission and improves the network lifetime. Several techniques have been presented for efficient data collection in WSN, that concern about the enhancing network lifetime. The Figure 1 represents Data Aggregation process [4].

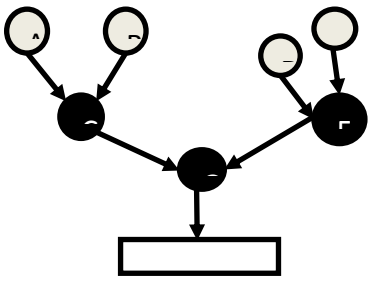


Fig. 1 Data Aggregation

As a traditional approach in WSN, nodes send data individually when the base station demands for network. But in Data aggregation approach, the aggregator, a special node is used to collect data from its neighboring stations, add them and forward that combined data to base station i.e sink node in multi-hop manner. The main aspect of data aggregation is to collect and aggregate data in an energy efficient manner by which the life span of network can be increased. Aggregating data is a process of comprising the transmitted packet, in the sense the packet should have only necessary information and no redundant data.

While designing a data aggregation algorithms some of the additional requirements to considered such as, they should take care of energy capabilities of sensor devices, energy resources and computational capabilities. And also the topology of the network to considered.

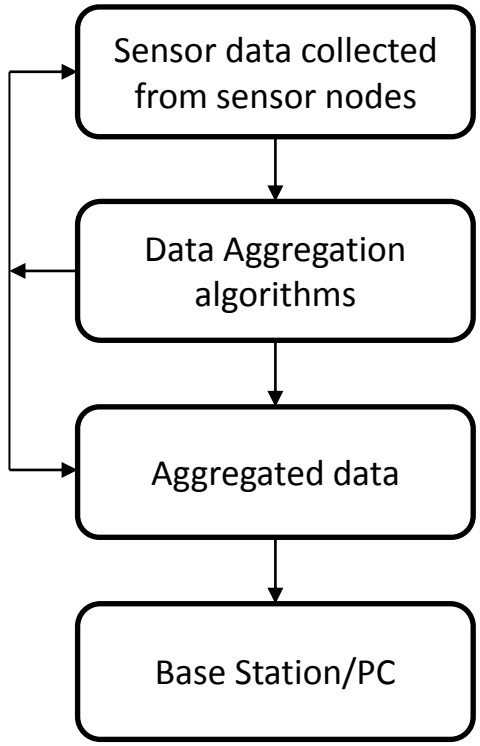
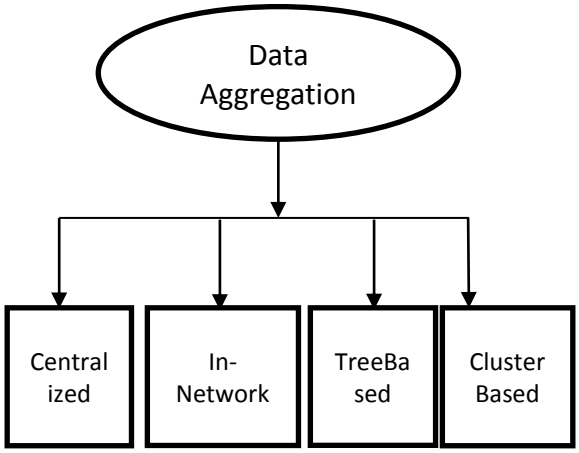


Fig. 2 shows the general architecture of data aggregation.

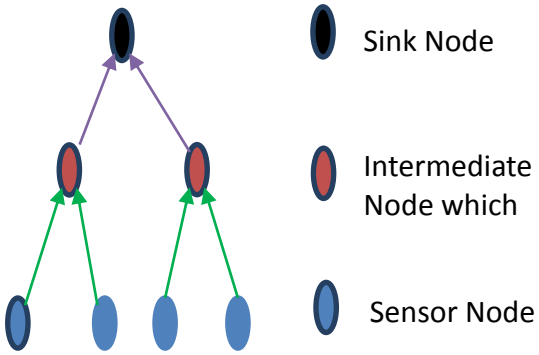


2.1 Data Aggregation Strategies

There are four strategies for aggregation some of them are listed below [5]:

Centralized Approach:

In this approach each sensor node sends its sensed data to a central node (base station) via the shortest possible route. All the sensor nodes simply sends the data packets to a node, which is the powerful among all other nodes. This node is called aggregator node or header node. This node aggregates the data coming from other nodes and the resultant data will be sent as a single packet.



In-Network Approach:

In network aggregation is a global approach for gathering and processing the data at intermediate nodes and routing the information through a multi-hop network. The main of this approach is to reducing power consumption.

There are two types of in-network aggregation:

1. With Size Reduction: Here the size of the packet to be transmitted to the sink node is reduced by combining and compressing the data packets received by sensor node from its neighbors.
2. Without Size reduction: Here, without processing the value of data the packets from the different neighboring nodes are merged into a single packet.

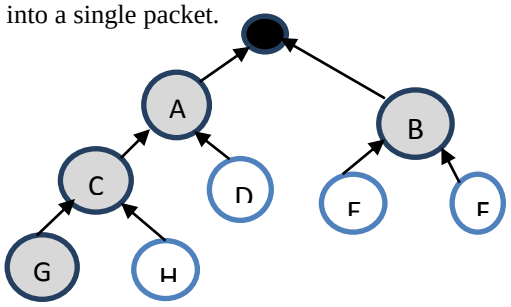


Fig.4: In-Network Aggregation

Tree Based Approach:

In this approach, a Data Aggregation Tree (DAT) is framed and here for each data transmission a minimum spanning tree is constructed. Each node in a network has a parent-child relationship in which the data is forwarded in a bottom-up approach. The data starts flowing from leaf nodes to the sink

node and the aggregation of the data is done by parent nodes in the network.

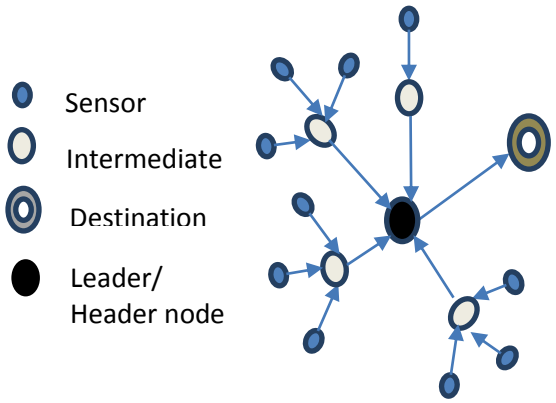


Fig. 5: Tree Based Data Aggregation

Cluster Based Approach

Here the whole network is split into several clusters. Each cluster is consisting of many sensor nodes. Cluster head is selected among the sensor nodes within a cluster. The aggregator role is performed by the Cluster head which aggregates the data received and send to the sink. By this approach the bandwidth overhead is minimized as total number of packets to be transmitted are less. Several clusters based approaches for data collection have been proposed for WSN.

Clustering reduces direct transmission to the base station by in network data aggregation as well as decreases energy consumption by reducing the transmitting distance. Better aggregation for large number of nodes is provided by Hierarchical Clustering [6, 7, 8, 9].

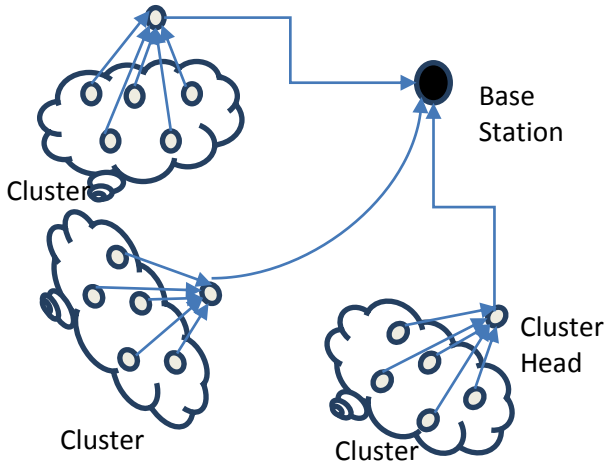


Fig.6: Cluster Based Data Aggregation

III. Cluster Based Protocols

LEACH(Low Energy Adaptive Clustering Hierarchy)-This protocol is a typical representation of hierarchical routing protocol. It is a self-adaptive and self-organized. Leach protocol uses round as unit, each round is made up of cluster set-up stage and steady state storage for the purpose of reducing unnecessary energy costs. The steady-state phase duration is usually much longer than set-up phase duration. However, the first phase is more important, in which sensor nodes are allowed to elect themselves as cluster-heads randomly, and then divided into clusters. Each node that becomes the cluster head (CH) will create a TDMA schedule for the sensor nodes within the cluster. That allows the radio components of each non CH-node to be turned off all times except during their transmit time. The job of the cluster-head is to collect data from their surrounding nodes and pass it on to the base station. LEACH is dynamic because the job of cluster-head rotates.

Some of the variants of LEACH for Homogeneous single hop network are:

i. LEACH_C (Low energy adaptive clustering hierarchy-Centralized)- Here the base station initiates centralized algorithm to elect the CHs according to their location info. It forms better balanced clusters. However it wastes energy to attain global information

ii. LEACH-F proposed is an efficient clustering technique based on LEACH protocol with clusters that are formed only once and then become fixed. In LEACH-F protocol, the cluster head role rotates among the nodes within the cluster. The rotation schedule of the future CH in the cluster is coordinated by the sink. In this way, the energy consumption is balanced between the sensors and the next CH is selected based on the status of each node.

iii. P-LEACH(Partition leach):It balances the wasted energy among sensors but with increase setup-phase. It achieved data transmission efficiency by making all clusters circle shaped; saved energy sensors by partitioning a cluster into four independent areas, thus distributing the roles of a managing sensor into four gate nodes; and could efficiently track mobile sink using distances between four gate nodes and the mobile sink.

iv. S-LEACH (Solar aware LEACH): Solar power improves network lifetime. It is used for both centralized and distributed CH selection algorithm.

v. E-LEACH: (Energy Efficient LEACH)[11]- E-LEACH is an energy efficient hierarchical cluster based protocol that distributes energy load evenly among all sensor units in the network. It is characterized by
   1) CH selection based on average energyconsumption and residual energy of nodes. Multi-hop inter-cluster communication that chooses a multi-hop path with minimal communication cost from each node to the BS.
   2) Direct communication by nodes near to the BS.

vi. RRCH (Round Robin Cluster head selection Algorithm)-The RRCH attain high energy efficiency by a single set-up process. It is as ingle setup process and minimizes energy consumption yet it cause extra overhead.

vii. CLUDDA: Clustered Diffusion with Dynamic Data Aggregation is a data aggregation based protocol that uses in-network processingto eliminate redundant transmission. It alleviates the flooding problem with Direct Diffusion (DD) by combining it with clustering. It also achieves dynamic dataaggregation points, since the location of the sink which initiates the interest can be changed and in turn new CHs and gateways that are closer to the sink will perform data aggregation. This improves the system performance through even distribution of energy consumption.

viii. WST-LEACH(Weighted Spanning Tree forLEACHthe selection of CHs depends on threeweighted parameters that optimize thetransmission path which in turn reduce powerdissipation that results in increasing networklifetime.

ix. EBC(Energy Balanced Clustering)-Here the re-clustering decisions are based on the traffic load processed by the CH in particular round instead of time scheduling, and this ensures that new CH selecting is done only when it is really required. Furthermore, it saves the energy resources spent during unnecessary re-clustering stages. EBC protocol balances energy and increases the overall lifetime of the network. But it causes extra overhead.

x. LEACH-SC (LEACH Selective Cluster)-It outperforms LEACH in terms of energy consumption and network lifetime by using location information.

Homogeneous Multi Hop:

i. M-LEACH (Multi-hop LEACH)- It is suitable for large size network but it suffers from hotspot and scalability. It improves LEACH protocol by allowing sensor nodes intra and inter-clusters to use multi hop communication in order to increase the energy efficiency of WSN.

ii. LEACH-L: It is an improved multi-hop routing protocol. It switches one-hop LEACH protocol to multi-hop transmission way according to the distance between CHs and BS or sink. If CH is close to the BS, single-hop strategy is adopted; otherwise the next hop toward the BS is selected based on its residual energy and distance to BS. LEACH-L protocol can balance network load, and reduce energy consumption of the sensors in different areas in addition to prolong the lifetime of WSNs.

iii. TL-LEACH(Two level LEACH)- This algorithm reduces the energy consumption by allocating the energy load among the sensors in dense networks. However it is not suitable for densely deployed network.

iv. MS-LEACH(Combination of multi-hop and single-hop)-It combines between multi-hop and single-hop Transmission modes to reduce energy consumption and prolong the lifetime of WSNs.

Heterogeneous Single Hop:

i. EECHE(Energy-Efficient Cluster Head Election Protocol)- It is a heterogeneous protocol that is based on three types of sensor networks. Each type is equipped with different energy resources, and the election of the CH is weighted based on the initial energy of a node. EECHE was developed for small sized wireless networks and it is based on single hop delivery from CHs to base station.

ii. NEAP(Novel Energy Adaptive protocol)-is based on LEACH and optimizes sensor nodes for the characteristics of heterogeneous WSN. However, the reliability of the network relies on the reliability of CH nodes. It is developed for monitoring environment remotely, and it is assumed that all sensor nodes have enough power to reach the base station.

Heterogeneous Multi-hop protocols

i. SEP (Stable election protocol) -It is developed to prolong the time interval before the first node dies. In SEP two levels of nodes are included: the advanced and normal nodes according to their initial energy.SEP does not require global knowledge of energy at every election round. It works with large- and small-scale networks, and does not require prior distribution of the sensor nodes. SEP is scalable since there is no need for position knowledge.

ii. HEED(Hybrid Energy-Efficient Distributed Clustering)- It operates in a multi-hop inter-clusterwireless sensor networks. It improves LEACH protocol by selecting periodically CHs based on combination of residual energy of each node and node's neighbor degree in order to achieve power balancing and increase the network scalability and lifetime.

iii. LEACH-HPR: Here CH selects the top stronger node as assistant node to balance the energy consumption. It select the intermediate nodes based on their residual energy. It uses the minimum spanning tree to construct an inter-cluster routing.

IV.RELATED WORK ON DATA AGGREGATION

1.In [12] author proposed an energy balanced and efficient data aggregation scheme for WSNs, called designated path(DP) scheme. DP scheme determined a set of paths and run them in round-robin fashion so that all the nodes can participate in the workload of gathering data and transferring the data to the sink. Nut the dissipated energy was increased.

2. In [13] author proposed an energy-efficient, secure, highly accurate, and scalable scheme for data aggregation (EESSDA). The main idea of EESSDA is that secure data aggregation is achieved by establishing secure channel and slicing technology. The EESSDA scheme does not need encryption and decryption operations during the data aggregation, which saves energy and obtain high accuracy of aggregation results. Meanwhile, in EESSDA scheme, the advanced deployment of shared information between nodes is not required, making the networks with good scalability.

3. In paper [14] the author developed a new energy efficient routing protocol called An energy efficient reliable routing protocol for wireless sensor networks(WSN) using data aggregation technique, Data aggregation is basically used to collect and aggregate data in an energy efficient manner so that network lifetime is improved.

4. In [15] the author applied a slice based energy model, and divided the energy balanced data collection problem into inter-and intra-slice energy balance problems.

5. Using a mobile element in current research to collect and carry data mechanically from a sensor network has many benefits over a multi-hop routing. In [16], the author proposes a data gathering mechanism for WSN by presenting the mobility in the network. A mobile collector namely M-collector could be a robot or be a vehicleequipped with a powerful battery and a transceiver, performed as a mobile base station and collects the data when moving through the area. The M-collector gathers the data periodically from the static sink, polls the every sensors when traversing its communication range and lastly transmits the data to the static sink.

6. In [17], the author proposed a polling-based mobile collection approach for WSN. A bounded relay hop mobile data collection (BRH-MDC) has been formulated as an optimization problem. A small number of sensors have been chosen as a polling points, which locally buffers the aggregated data and the data are uploaded to the mobile collected while it arrives. While the sensor allied with these points, it is assured that any data relay is bounded within a provided number of hops. Additionally, two algorithms are presented for choosing the polling points among sensors.

7. In [19], the author have proposed a compressed data aggregation scheme that exploits compressed sensing (CS) technique to conquer both recovery fidelity and energy efficiency in WSNs with arbitrary topology.

8. In [20] the author introduced data density correlation degree and the data density correlation degree (DDCD) clustering method. With the DDCD clustering method, the sensor nodes that have high correlation are divided into the same cluster, allowing more accurate aggregated data can be obtained in cluster-based data aggregation networks produced by the DDCD clustering method. Also, the amount of data conveyed to the sink node can decrease.

9. In [21], the author proposed a new data aggregation algorithm named ERDL (Efficient and Real time algorithm based on Dynamic message List). ERDL works based on network layer of WSNs, and a dynamic list will be created in filtering node to store history messages ever relayed by this node. All messages in WSNs will be judged whether reduplicated or not according to the contents in list. In ERDL the filtering efficiency is improved, and the real time performance of transmitting is also ensured.

10. In [22], the author proposed a Velocity Energy-efficient and Link-aware Cluster-Tree (VELCT) scheme for data collection in WSNs which would effectively mitigate the problems of coverage distance, mobility, delay, traffic, tree intensity, and end-to-end connection. The proposed VELCT constructs the Data Collection Tree (DCT) based on the cluster head location. The data collection node in the DCT does not participate in sensing on this particular round, however, it simply collects the data packet from the cluster head and delivers it to the sink. The designed VELCT scheme minimizes the energy exploitation, reduces the end-to-end delay and traffic in cluster head in WSNs by effective usage of the DCT. The strength of the VELCT algorithm is to construct a simple tree structure, thereby reducing the energy consumption of the cluster head and avoids frequent cluster formation. It also maintains the cluster for a considerable amount of time.

11. In [23] the author proposed data fusion approach for resource efficiency in large WSN. Data fusion is used to determine a reduced node set to be active in the network, resulting in reduction of network resource consumptions.

12. In [24], author proposed to utilize mobility for joint energy replenishment and data gathering. A multi-functional mobile entity, called SenCar was employed, not only as a mobile data collector roaming over the field to gather data via short-range communication but also as an energy transporter that charges static sensors on its migration tour via wireless energy transmissions.

13. In [25], author Shiliang Xiao et al. have exploited the tradeoff between data quality and energy consumption in order to improve the data aggregation precision in case of heterogeneous per-node energy constraints.

14. In [26], author presented a data aggregation architecture model integrating a multi-resolution hierarchical structure with CS to further optimize the amount of data transmitted.

15. In [27], author SongtaGuo et al proposed a framework of joint wireless energy replenishment and anchor-point based mobile data gathering in WSNs by view of various sources of energy consumption and time-varying nature of energy replenishment.

V. CONCLUSION

WSN is an energy constrained network. Since most of the energy has depleted for transmitting and receiving data, so the data aggregation becomes essential in the network. Data aggregation or data gathering helps to eliminate duplicate data

transmission in WSN. It has attracted a lot of attention in the recent time. In this paper, we have summarized some of research results on data gathering and data routing in WSN. Resilience to setup overhead, routine link breaks, Setup, Scalability, mobility of nodes, power saving method, and Timing strategy, data aggregation protocols with cluster approach perform well compared with other protocols and can build energy efficient WSN with these protocols. We surveyed recent proposed clustering protocols for WSNs and classified these into four categories depending on the network topology as well as the hop communication i.e homogeneous, heterogeneous, single and multi-hop. A proper study of the relation between energy efficiency and network lifetime is a channel for the future research.

REFERENCES

[1] Sumit Chaudhary et al, "Energy Efficient Techniques for Data Aggregation and Collection in WSN ", *International Journal of Computer Science, Engineering and Applications*, Vol. 2, No. 4, August 2012.

[2] Mousam Dagar and Shilpa Mahajan, "Data Aggregation in Wireless Sensor Network: A Survey", *International Journal of Information and Computation Technology*, ISSN 0974-2239 Vol. 3, No. 3, 2013, pp. 167-174

[3] V.Akhila, T.Sheela, G.Adiline Macriga, "Efficient Packet Scheduling Technique for Data Merging in Wireless Sensor Networks", CHINA Communications, April 2017.

[4] M.Suraj, B.Raja, T.Venkataraman, 'Secure Data Aggregation In Wireless Sensor Network', International journalof Computer Science Trends and Technology (IJCST)-volume 4 issue 2, Mar-Apr 2016.

[5] Binkal S Ahir,Rohan Parmar, Bintu Kadhiwala,' Energy efficient Clustering Algorithm for Data Aggregation in WSN', International Conference on Green Computing and Internet of Things(ICGCIoT),2015.

[6] PeimanGhaffariyan, "An Effective Data Aggregation Mechanism for Wireless Sensor Networks", 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM),pages 1-4,September 2010.

[7] Yaeghoobi, K.S.B.; Tyagi, S.S.; Soni, M.K.; Ebadati E, O.M., "SAERP: An energy efficiency Real-time Routing protocol in WSNs," in Optimization, Reliabilty, and Information Technology (ICROIT), 2014 International Conference on , vol., no., pp.249-254, 6-8 Feb. 2014.

[8] Sabri, Alia, and Khalil Al-Shqeerat. "Hierarchical Cluster-Based Routing Protocols for Wireless Sensor Networks–A Survey." IJCSI International Journal of Computer Science Issues 11.1 (2014).

[9] M. MehdiAfsar, Mohammad-H and Tayarani-N, "Clustering in sensor networks: A literature survey", Journal of Network and Computer Applications 46, pp 198--226, 2014.

[10] Alia Sabri , Khalil Al-Shqeerat,' Hierarchical Cluster-Based Routing Protocols for Wireless Sensor Networks – A Survey', IJCSI International Journal of Computer Science Issues, Vol. 11, Issue 1, No 2, January 2014 ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784

[11] Anuradha Rai, Suman Deswal, Dr. Parvinder Singh,' An Energy-Efficient E-LEACH Protocol for Wireless Sensor Networks', ISSN 2321 3361 © 2016 IJESC, Volume 6,Issue No.7

[12] RabindraBista, Yong-Ki Kim, and Jae-Woo Chang, "A New Approach for Energy-Balanced Data Aggregation in Wireless Sensor Networks", IEEE Ninth International Conference on Computer and Information Technology, China ,Volume 2,October 2009.

[13] PeimanGhaffariyan, "An Effective Data Aggregation Mechanism for Wireless Sensor Networks", 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM),pages 1-4,September 2010.

[14] Basavaraj S. Mathapati, Siddarama. R. Patil and V. D. Mytri, "Energy Efficient Reliable Data Aggregation Technique for Wireless Sensor Networks", International Conference on Computing Sciences, Phagwara , Pages 153-158 ,14-15 September 2012.

[15] Ning Jin, Kaiji Chen and Tao Gu, "Energy Balanced Data Collection in Wireless Sensor Networks", 20th IEEE International Conference on Network Protocols (ICNP), pages 1-10,October2012.

[16] M. Ma, Y. Yang, and M. Zhao, "Tour planning for mobile data gathering mechanisms in wireless sensor networks," IEEE Transactions on Vehicular technology, Vol. 62, No. 4, 2013,pp. 1472–1483.

[17] M. Zhao and Y. Yang, "Bounded relay hop mobile data gathering in wireless sensor networks," IEEE Transactions on Computing, Vol .61, No. 2, 2012,pp. 265–271.

[18] Liu Xiang, Jun Luo, and Catherine Rosenberg, "Compressed Data Aggregation: Energy-Efficient and High-Fidelity Data Collection", IEEE/ACM transactions on Networking, vol. 21, no. 6, pp 1722-1735,December 2013.

[19] Fei Yuan, Yiju Zhan, and Yonghua Wang, "Data Density Correlation Degree Clustering Method for Data Aggregation in WSN", IEEE Sensors Journal, Vol. 14, No. 4, Pages 1089-1098,April 2014.

[20] TaoDu,ShouningQu,KaiqiangLiu, JinwenXb. YinghuaCao, 'An Efficient Data Aggregation Algorithm for WSNs Based on Dynamic Message List' Procedia Computer ScienceVolume 83, 2016, Pages 98-106,Elsevier.

[21] R. Velmani, B. Kaarthick,' An Efficient Cluster-Tree Based Data Collection Scheme for Large Mobile Wireless Sensor Networks', IEEE Sensors Journal (Volume: 15, Issue: 4, April 2015).

[22] MohammadrezaSoltani, Michael Hempel, and Hamid Sharif, "Data Fusion Utilization for optimizing Large-Scale Wireless Sensor Networks", IEEE ICC 2014 - Ad-hoc and Sensor Networking Symposium, pages 367-372,June2014.

[23] Miao Zhao, Ji Li, and Yuanyuan Yang, "A Framework of Joint Mobile Energy Replenishment and Data Gathering in Wireless Rechargeable Sensor Networks", IEEE Transactions On Mobile Computing,Pages 2689-2705, Vol. 13, No. 12, December 2014.

[24] Shiliang Xiao, Baoqing Li, Xiaobing Yuan, "Maximizing precision for energy-efficient data aggregation in wireless sensor networks with lossy links", Ad Hoc Networks,Elsevier, Vol 26,pages 103–113 ,March 2015.

[25] Xi Xu, Rashid Ansari and AshfaqKhokhar, "Power-efficient Hierarchical Data Aggregation using Compressive Sensing in WSNs", IEEE ICC - Ad-hoc and Sensor Networking Symposium, pages 1769-1773,2013.

[26] SongtaoGuo, Cong Wang and Yuanyuan Yang, "Mobile Data Gathering with Wireless Energy Replenishment in Rechargeable Sensor Networks", Proceedings IEEE INFOCOM, pages 1932-1940,2013.

[27] M. MehdiAfsar, Mohammad-H and Tayarani-N, "Clustering in sensor networks: A literature survey", Journal of Network and Computer Applications 46, pp 198--226, 2014.